This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.





1 Publication number:

0 499 681 A1

12

EUROPEAN PATENT APPLICATION

2) Application number: 91106110.9

(s) Int. CI.5: **C12N** 15/30, A61K 39/118, G01N 33/569, C07K 15/00

2 Date of filing: 17.04.91

Priority: IT MI91A000314.

(3) Priority: 07.02.91 IT 31491

© Date of publication of application: 26.08.92 Bulletin 92/35

Designated Contracting States:
 AT BE CH DE DK ES FR GB GR IT LI LU NL SE

7) Applicant: SCLAVO S.p.A. Vla Florentina 1 I-53100 Slena(IT)

2 Inventor: Ratti, Giulio Strada di Vico Alto, 6 I-53100 Siena(IT) Inventor: Comanducci, Maurizio Via Montanini, 54 I-53100 Siena(IT) Inventor: Tecce, Mario F. Via Mino Celsi, 12

I-53010 Costalfino (Siena)(IT) Inventor: Giuliani, Marzia M. Via Garibaldi, 58

I-53034 Colle val d'Elsa (Siena)(IT)

(A) Representative: Moretti, Giorgio et al Notarbartolo & Gervasi s.r.l. Viale Bianca Maria, 33 I-20122 Milano(IT)

- PCTD plasmid isolated form chlamydia trachomatis serotype D, its genes and proteins encoded by them; recombinant plasmids for the expression of said genes in heterologous systems as fused recombinant proteins, preparation of said recombinant proteins and their use in the formulation of vaccins and/or diagnostics.
- A plasmid isolated from Clamydia trachomatis is described, which comprises 8 genes encoding proteins useful in the formulation of vaccines or diagnostic test for determining the bacterium or specific antibodies generated during C. trachomatis infections; in particular the recombinant fusion MS2-pgp3D protein is described comprising polypeptidic sequences encoded by pCT and immunogenic in the course of infections in man. A method for preparing said protein in E.coli further described.

Invention Field

This invention refers to the pCTD plasmid isolated from Chlamydia trachomatis serotype D, cloned and sequenced and to the genes present in said plasmid, to the proteins expressed by said genes, to the expression vectors containing said genes and to the microrganisms transformed by said vectors. The invention further refers to the process for the preparation of genes and of said vectors and to the use of said proteins as antigens for the preparation of polyclonal and monoclonal antibodies apt to recognize Chlamydia trachomatis and hence useful for the preparation of vaccines capable of imparting a protective immunity against infections caused by Chlamydia trachomatis and pathologic conditions deriving from said infections and for the development of diagnostic methods for the search of specific antibodies produced following C.trachomatis infections.

Prior art

15

20

40

Chlamydias are gram-negative bacteria, obligate intracellular parasites of eukariotic cells. Chlamydias show an extracellular infective and metabolically practically inert form, called elemental body (EB), and intracellular replicative forms called reticular bodies (RB).

The reticular bodies, after multiplication by binary fission, are transformed into elemental bodies which come out of the host cell and infect new cells.

The masses or mini-colonies of reticular and elemtal bodies inside an infected cell constitute the characteristic "inclusions" visible at the optical microscope.

Chlamydia trachomatis (C.trachomatis or CT), a bacterial species pathogenic to man, is the etiological agent of venereal lymphogranuloma (VLG), of various inflammatory patologies of the genital male and female apparatus and of trachoma, a chronic disease which affects 500 million people and can lead to blindness.

In the technical literature ca. 15 CT serotypes pathogenic to man were described and divided in two groups which differ both as to virulence and tissular tropism.

Twelve serotypes of the trachoma group (biovar) are identified as A to K and infect, in general, epithelial tissues, such as the ocular (trachoma) and uro-genital (cervicitis and urethritis) mucous membranes; and show a low virulence.

The venereal lymphogranuloma (VLG) serotypes (L₁, L₂ and L₃) cause instead an infection of the reticulo-endothelial tissue, mainly of the inguinal and femoral lymphonodi, and are highly invasive.

Urethritis and cervicitis induced by CT (A to K serotypes) when not precociously diagnosed and treated by adequate therapy, may led to a variety of chronic inflammations, such as, e.g., vaginitis, salpingities and pelvic inflammation which may resolve in sterility and extrauterine pregnancy.

Furthermore the new born from infected mothers may contract pulmonary and/or ocular infections during delivery.

For said reason it is necessary to possess adequate diagnostic methods for determining CT and formulating effective vaccines against said bacterium.

As known, factors which determine the bacterial virulence are often encoded by genes present on plasmids.

In the literature, the presence is reported, in all 15 serotypes and in the clinical isolates examined up to now, of a plasmid of ca. 7.5 Kb referred to in the present invention as pCT followed by the denomination of the bacterial serotype concerned. For example: pCTD for the plasmid isolated from serotype D, etc.

Up to now, however, no specific function or products encoded by it were associated with said plasmid.

Detailed description of the invention

A variant of the plasmid, corresponding to serotype D, was now isolated, indicated in what follows a pCTD, which comprises at least eight genes encoding for new proteins.

Figure 1a shows the nucleotidic sequence of said plasmid and 7 of the 8 protein structures expressed by said sequence. The eighth protein structure, encoded on the DNA chain complemental to the one of Fig. 1a, is shown in Fig. 1b.

Object of the present invention are thus: the cloned and sequenced pCTD plasmid, the nucleotide sequences encoding for the above named proteins, the expression vectors containing one of said sequences or fragments thereof.

Further object of the present invention are the pCTD proteins or fragments of them having immunogenic properties.

Still another object of the present invention are the fusion polypeptides comprising one of said proteins or its fragments suitable as antigens.

The present invention further refers to the preparation of said proteins and of their fragments possessing immunogenic activity or of fused polypeptides comprising said proteins.

Said proteins, their fragments or fusion polypeptides comprising said proteins or their fragments, according to the invention may be employed to determine the CT produced infections in biological samples.

Said proteins, their fragments or fusion polypeptides comprising the protein or its fragments may further be employed, according to the invention, as antigens useful in the formulation of vaccines against infections due to CT.

According to the invention, said proteins, their fragments or fusion polypeptides may be used furthermore as antigens for the preparation of poly- or mono-clonal antibodies to be used in diagnostics. In particular, the present invention relates to the pgp 3D protein encoded by the gene of the pCTD plasmid identified as ORF3D having the nucleotide sequence reported in Fig. 2, and characterized by a molecular weight of 27,802 and by the aminoacid sequence reported in Fig. 2.

According to the present invention, plasmid pCTD is obtained from the C.trachomatis GO/86 strain isolated from the urethra of a patient with non-gonococcic urethritis, and successively identified as serotype D by the immunofluorescence method described by Wang, S.P. and Grayston, J.T. [Am. J. Ophtalmol. 70; 367-374 (1970)]. The ORF3D gene may be isolated from the pCTD plasmid employing one of the known methods such as, e.g., the in vitro amplification method [Saiki, A.K. et al. Science, 239:487-491 (1988)] using as primers synthetic oligonucleotides having a primary structure suitably derived from the sequence data shown in Figs. 1a and 1b. The thus emplified gene is then cloned in a vector placing it under the control of sequences regulating its expression.

One can similarly proceed for the other seven genes the nucleotide sequences of which are reported in Figs. 1a and 1b.

The proteins encoded by said genes are represented by the aminoacid sequences also reported in Figs. 1a and 1b.

Vectors suitable for the ends of the present invention may be plasmids with expression in host cells selected among the ones known and available commercially or at authorized collection centers.

The cells transformed by said vectors are then cultivated in a suitable culture medium in the presence of carbon-, nitrogen- and mineral salts sources, possibly in induction conditions, at a temperature and time period selected in order to obtain the production of the desired protein.

Said protein, obtainable also as fused polypeptide, constituted by a polypeptide produced by the vector fused with the protein itself, is then separated and purified from the culture medium or from the cell lysate.

According to one embodiment of the present invention, the ORF3D gene is cloned in the plasmidic E.coli pEX34a vector, a derivative of pEX29 and pEX31 described by Strebel et al. [J.Virol., 57:983-991 (1986)], following the description by Nicosia et al. in Infect. Imm. 1987, Vol.55, 963-967.

The results show the presence in the bacterial extracts of a polypeptide, indicated as MS2-pgp3D, the sequence of which is shown in Fig. 3, with a mol. weight of ca. 39 Kd, consisting i.e. of a RNA-polymerase fragment of bacteriofage MS2, produced by the expression system of ca. 11 Kd and by the protein encoded by the ORF3D gene of ca. 28 Kd.

Said polypeptide employed as antigen in a Western-Blot assay, or in immunologic assays, is recognized by antibodies present in the serum of patients with CT infection and may further be employed for the production, in laboratory animals, of mono- and poly-clonal antibodies which recognize the - and react with the corresponding pgp3 protein, in all its variants, of C.trachomatis.

In accordance with the present invention the pCTD and p03/60/MCI plasmids were deposited as ATCC N° 68314 and ATCC N° 68315 respectively.

The experimental examples that follow are illustrative and non limitative of the invention.

EXAMPLE 1

45

50

10

15

Isolation of the pCTD plasmid from C.trachomatis GO/86

C.trachomatis cells were isolated following known techniques from the urethra of a patient with non-genococcic urethritis. The strain, identified as serotype D by the micro-immunofluorescence technique described by Wang, S.P. and Grayston, J.T. [(1970), Am. J. Ophtalmol., 70: 367-374] is designated as GO/86.

The elemental bodies of said strain are then purified as described by Cevenini R. et al. [(1988), FEMS Microbiol. Letters, 56:41-46] on renografin^R density discontinuous gradients (E.R. Squibb & Sons, Princeton,

N.J.) according to what reported by Caldwell H.D. et al. [(1988) Infect. Immun. 31:1161-1176].

After purification, the elemental bodies (ca. 1.5 mg proteins) are lysated by incubation in 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2mM EDTA, 0.6% SDS and 100 mg/ml K Proteinase (Boehringer) at 37 °C for 3 hrs. The total nucleic acids are then extracted with phenol/chloroform, precipitated with ethanol, treated with pancreatic RNAse (250 ng/µl final concentration), further precipitated with ethanol and re-suspended in 800 µl water (365 ng/µl of DNA).

A 10 μ I aliquot of said solution is then treated with 30 units (U) of BamHI restriction enzyme (Boehringer) at 37 °C for 2 hrs in 20 μ I (final volume) of a digestion mixture suggested by the supplier. 3 μ I of the resulting digestion mixture are ligated to 100 ng plasmidic pUC8 DNA previously digested with BamHI and dephosphorilated with calf gut phosphatase. The ligase reaction is effected overnight in 20 μ I buffer containing 9 U T4 DNA ligase (Boehringer) at 18 °C.

The ligation mixture is then employed to transform HB101 E.coli cells made competent by a treatment with CaCl₂ as described by Mandel and Higa [(1970) J. Mol. Biol: 53, 54]. The transformants are selected on LB agar Medium (DIFCO) with addition of 100 μg/ml ampicillin, at 37°C overnight.

The positive clones (ampicillin resistant) (Amp^R) containing, that is, the recombinant pUC8 plasmid are transferred onto Hybond-N membranes (Amersham) and sorted by hybridization with three ³²P marked oligonucleotides having the following nucleotidic sequences:

- 1) 5'ATGGGTAAAGGGATTTTATC3'
- 2) 5'CTATATTAGAGCCATCTTC3'
- 3) 5'TCAAAGCGCTTGCACGAAG3'

The above reported oligonucleotides are synthesized by means of an automatic synthesizer (Applied Biosystem Inc. Mod. 380A) following the methods and employing the reagents recommended by the

Four of the six plasmids isolated from the clones found positive at the hybridization, analyzed by electrophoresis on agarose 1% gel before and after digestion with BamHl are found to consist of the pUC8 plasmid nucleotidic sequence and of a nucleotidic insert of ca. 7.5 kilobases corresponding to the isolated C.trachomatis GO/86 plasmid.

The nucleotidic sequences of said insert is determined according to the method of Sanger F. [(1977) PNAS USA 74:5463-5467] utilizing a series of suitable primers. The sequencing reactions are performed on double helix DNA employing the Sequenase Kit (U.S. Biochemical Co. Cleveland, Ohio) as recommended by the firm.

The nucleotidic sequences of the ca. 7.5 kilobases plasmid named pCTD are reported in Figs. 1a and 1b. The recombinant plasmid containing said insert is indicated as pUC8-pCTD.

EXAMPLE 2

Cloning of the DNA ORF3D segment of plasmid pCTD1D

The DNA fragment denoted as ORF3D(Fig. 2) of 792 bp is obtained through in vitro amplification according to the technique known as Polymerase Chain Reaction (PCR) described by Saiki A.K. et al. [-(1988) Science 239:487-491].

The amplification is effected utilizing ca. 10 ng of the pUC8-pCTD plasmid and employing as primers two synthetic oligonucleotides (ORF31) and (ORF3dx) having respectively the following nucleotide sequences:

ΔŊ

15

20

- 5'CAGGGATCCATGGGAAATTCTGGTTTTT3'

BamHI

5

10

15

- 5'CCCCTGCAGTTAAGCGTTTGTTTGAGGT3'

Pst I

Said oligonucleotides are complemental to ORF3 regions with the addition to the respective 5' terminals of a nucleotide sequence comprising the action site of a restriction enzyme selected among the ones present in the pEX34A vector (Strebel K. et al. [(1986) J. Virol.57: 983-991] utilized for the successive cloning. In particular, the site selected for ORF31 is the one for the BamHI enzyme, while for ORF3dx is the one of the PstI enzyme.

The amplification reaction is performed employing the reagents contained in the "Geneamp" Kit (Perkin Elmer-Cetus). 25 amplification cycles are effected. Each amplification cycle consists in heating the reaction mixture to 94 °C for one minute, to 50 °C for one minute and finally to 72 °C for one minute.

At the end of the amplification reaction the mixture is extracted, in succession, with an equal volume of phenol and of a chloroform-isoamyl alcohol mixture (24:1 v/v) and then submitted to forced dialysis by means of Centricon^R cartridges following the producer's (Amicon) instructions.

The DNA is then precipitated by adding to the obtained solution sodium acetate 3 M, pH 5.5 (1/10 of the volume) and cold (-20 $^{\circ}$ C) ethanol (3 vols.). The DNA precipitate is dissolved in 44 μ l water. To the solution, 5 μ l H buffer (Boehringer) and 1 μ l PSTI restriction enzyme (20 units/ μ l) are added and the DNA is digested at 37 $^{\circ}$ C for 2 hours.

The digestion mixture is then extracted with phenol, chloroform/isoamyl alcohol and then the DNA is precipitated with ethanol (-20 °C). The precipitate, separated by centrifugation, is suspended again in 44 µl water and then digested with 20 U BamHl in 5 µl of B buffer (Boehringer) at 37 °C for 2 hours. The digestion mixture is extracted with phenol, chloroform/isoamyl alcohol and dialyzed by Centricon^R cartridge.

At the same time, 10 µg of the pEX34A plasmidic vector are digested with the Pstl and BamHI restriction enzymes as reported supra. The vector is dephosphorylated with alkaline phosphatase, extracted with phenol and chloroform/isoamyl alcohol, precipitated with ethanol (-20 °C) and re-suspended in 50 µl water.

1 μI (100 ng) of the vector and 2 μI (200 ng) of the amplified ORF3D segment are then ligated in 2 μI ligase buffer to which 2 μI ATP r, 1μI T4 DNA ligase (9 units/μI) are added, adding water to a total volume of 20 μI. The ligase reaction is performed at 15°C overnight. The ligase mixture is employed to transform 200 μI of a suspension of E.coli competent cells (K12-ΔH1-Δ trp) [Remaut E. et al. (1983), Gene 22:103-113]. After treatment at 30°C for 5 minutes, to the cell suspension 800 μI LB medium are added, followed by incubation at 30°C for 1 hour. Aliquots of the cell suspension (10 μI, 100 μI and 690 μI) are separately plated on plates of agarized (20 g/l) LB medium containing 100 μg/mg ampicillin and kept at 30°C overnight.

The obtained clones (Amp^R) are transferred to a nitrocellulose membrane on a LB agar plate with added ampicillin, grown at 30 °C overnight, and then tested for hydridization with three oligonucleotidic probes (UB35, UB36, UB18) terminally marked with ³²P having the following sequences:

45

50

- I) 5'-ATGGGTAAAGGGATTITATC3'
- II) 5'-CTATATTAGAGCCATCTTC3'
- III) 5'-TCAAAGCGCTTGCACGAAG3'

The hybridization test is performed according to known tecnique. From the colonies positive to hybridization the plasmids contained in them are prepared by minipreparation as described by Maniatis et al. (1982) and the ORF3D insert nucleotide sequence is controlled by known technique.

EXAMPLE 3

Expression of the MS2-gpg3 recombination protein

E.coli cells containing the pEX34 vector with the ORF3D insert are inoculated in duplicate in 10 ml LB medium with added 30 μg/ml ampicillin and cultivated at 30 °C overnight. The procedure described by Nicosia et al. [Inf. Imm. (1987) 55:963-967] is then followed, with the provision that one of two duplicates undergoes induction of the cloned gene by treatment at 42 °C, while the other does not. Two protein extracts are thus obtained, produced by the bacterium, in 7M urea buffered at pH 8, one of which corresponds to the induced cells, and the other, as a control, to the non-induced cells.

By analysis of the protein contents of both extracts by electrophoresis in SDS-polyacrylamide 15% gel according to known techniques, it is possible to deduct the presence of a protein species of 39,000 apparent mol.wt. which is present in a considerably greater amount in the induced extracts.

In the non-induced cell lysate no evidence of such a protein, but only the product of the vector alone, is found.

Said electrophoresis patterns may be analyzed by the Western Blot technique employing a monoclonal antibody (SCLAVO) specific for the 11 kd fragment generated by the pEX34 vector. In this way it is possible to demonstrate that the 39 kd band is a fusion protein containing said fragment.

EXAMPLE 4

Purification of MS2-pgp3 from E.coli K12\Delta H1\Delta trp extracts

The protein extract, from induced bacterial cells, re-suspended in 7M urea, is dialyzed for 15 hrs. at 4°C against a PBS buffer consisting of 0.4% KCI, 0.4% KH₂PO₄, 16% NaCI, 2.5% NaH₂PO₄.

During the dialysis a protein precipitate is obtained, which is separated by centrifuging and discarded. The surnatant is submitted to further purification by electrophoresis on preparative 12.5% acrylamide gels, and the protein band of 39,000 mol.wt. (MS2-pgp3D) is then extracted by electroelution from the gel.

The thus obtained MS2-pgp3 is precipitated by adding to the electroeluted solution 9 volumes of absolute acetone (-20°C). The protein precipitate is separated by centrifuging, re-suspended in 90% acetone, centrifuged as above, precipitated in 96% acetone and centrifuged again. The precipitate is brought to dryness in a nitrogen stream and re-suspended in 200 µI sterile PBS at a final concentration of approximately 1.5 µg/µI.

The advantage of the effected dialysis is the elimination, with this procedure, of some E.coli proteins, in particular some with a molecular weight equal or very near to the one of the desired recombinant product, which may present a considerable hinderance in the electrophoretic and/or chromatographic purification.

EXAMPLE 5

55

Production of polyclonal anti-MS2-pGPG3 antibodies

Utilizing the MS2-pgp3 protein, purified as in Example 4, 3 Balb/C 7-8 week old mice are immunized intraperitoneally. The immunization procedure comprises a first injection of 0.2 ml/mouse of an emulsion consisting of one part by vol. of the purified protein solution (1.5 µg/µml) and five parts of Freund complete adjuvant (FCA).

The thus inoculated protein amount is thus ca. 50 μ g/mouse. After 1 week the mice are immunized with the said same emulsion, followed by a 800 μ l Pristane injection. After 1 week from the second inoculation, the mice are intraperitoneally immunized with 0.2 ml of a solution similar to the first one. Finally, after two weeks from the third inoculation a booster immunization is effected. The thus induced antibodies are collected in the ascitic fluid formed after the above described treatment.

The anti MS2-pgp3 antibody titres show values comprised between 1:8000 and 1:10.000 evaluated by analysis with Western Blot containing the MS2-pgp3 protein.

The reactivity of said antibodies to the native antigen (pgp3) was evaluated according to the following methods:

- analysis with Western Blot containing total protein extracts of elemental purified CT bodies
- immunofluorescence on McCoy cells cultures infected with CT. The results of the above tests show that the anti MS2-pgp3 antibodies are able to reveal C.trachomatis inclusions in infected cells (see immunofluorescence test) and recognize a protein present in the bacterium protein extracts and having a mol.wt. of 28 kd, equivalent, that is, to the one of the protein encoded by ORF3D (see Western Blot test).

EXAMPLE 6

To the end of preparing monoclonal anti-MS2-pgp3 antibodies, the mice, immunized as above described, are sacrificed, the spleens extracted and utilized for the preparation of hybridomas operating according to the technique described by Davis L.G. [Basic methods in molecular biology - Elsevier Edit., New York (1986)]. The screening of the thus obtained hybridomas is performed as described for the polyclonal antibodies. In particular, a screening was performed with induced E.coli extracts (see Example 3) containing the MS2-pgp3 protein or the polypeptide encoded by the pEX34 vector alone; obviously, the clones were selected which produced antibodies reacting only with the recombinant product. With such pgp3-specific antibodies, results are obtained which are superimposable to the ones obtained with the above described polyclonal antibodies.

EXAMPLE 7

15

25

30

35

40

45

50

55

Serum samples from 20 patients with Chlamydia generated infections were collected. Said sera contained anti-Chlamydia antibodies with titres comprised between 128 and 512, as determined by immunofluorescence against single antigen (LGV2). 15 control sera not containing anti-Chlamydia antibodies were obtained from healty donors. Western Blots were prepared, as above described, containing the MS2-pgp3 protein. These were incubated with the sera under examination diluted 1:100 and successively with peroxidase marked rabbit (anti human IgG) immunoglobines. 16 of the 20 infected patients sera contained antibodies apt to react with MS2-pgp3. The 15 healthy control sera did not give any reaction with said protein.

Claims

1. pCTD plasmid isolated from Chlamydia trachomatis serotype D characterized by the following nucleotidic sequence:

	10	30	50	
5	ATATTCATATTCTGTTGCCAG	AAAAAACACCTTTAGGC	TATATTAGAGCCATCTTCTTTG	
	70	90	110	
	AAGCGTTGTCTTCTCGAGAAC	ATTTATCGTACGCAAAT.	ATCATCTTTGCGGTTGCGTGTC	
	130	150	170	
10	CTGTGACCTTCATTATGTCGC	AGTCTGAGCACCCTAGG	CGTTTGTACTCCGTCACAGCGG	
			230	
	190	210 	GCAGCTTGTAGTCCTGCTTGAG	
	TTGCTCGAAGCACGTGCGGG			
15	250	270	290	
	AGAACGTGCGGGCGATTTGC	TTAACCCCACCATTTTT	CCGGAGCGAGTTACGAAGACAA	
	310	330	350	
	AACCTCTTCGTTGACCGATG	PACTCTTGTAGAAAGTGC	ATAAACTTCTGAGGATAAGTTA	
20			410	
	370	390 PGACGGTTCTTAAGCTGG	GAGAAAGAAATGGTAGCTTGTT	
	430	450		
?5	GGAAACAAATCTGACTAATC	TCCAAGCTTAAGACTTCA	GAGGAGCGTTTACCTCCTTGGA	
-	490	510	530	
	GCATTGTCTGGGCGATCAAC	CANTCCCGGGCATTGATI	TTTTTTAGCTCTTTTAGGAAGG	,
	EEA	570	590	
30	550 ATGCTGTTTGCAAACTGTTC		PATTTCCCTGGTTTTAAAAAATG	,
	610	630 CAACCUUCCGCUATAGCC	650 GACTATTCCTTGAGTCATCCTGI	٠
	TTCGACTATTTCTTGTTTA	ONNOUT TOCUC IN THUCK	one in i control of the control of t	
×	670	690	710	
,,	TTAGGAATCTTGTTAAGGAA	ATATAGCTTGCTGCTCG	AACTTGTTTAGTACCTTCGGTCC	
	730	750	770	
	AAGAAGTCTTGGCAGAGGAA	ACTTTTTTAATCGCATC	taggattagattatgattt aaa j	L
••	790	810	830	
Ю			ACCARTCITITCTARAGACAAA	٨
	GGGAMAC ICI I COMO			
	850	870	890 TGATGCGGTCCAATGCATAATA	
	AAGATCCTCGATATGATCT/	CAAGTATGTTTGTTGAG	IGATGCGGTCCAATGCATAATA	•
15	910	930	950	
	CTTCGAATAAGGAGAAGCT	PTTCATGCGTTTCCAATA	GGATTCTTGGCGAATTTTTAAA	À
	970	990	1010	
	CTTCCTGATAAGACTTTTC		TTCTTGCTGCAAAGATAAAATC	C
50				
	1030	1050	1070 TAAAATCTCCTCATTAGTGAAA	T
	CTTTACCCATGAAATCCCT	CGTGATATAACCTATCCG	TAAAATGTCCTGATTAGTGAAA	•
	1090	1110	1130	_
55	AATCAGGŤTGTTAACAGGA	TAGCACGCTCGGTATTT	TTTATATAAACATGAAAACTCG ORF1 >> MetLysThrAr	T
			CORFI 33 RETLYSIDIAL	-

5	. 1150 TCCGAAATAGAAAATCGCATG	1170 CAAGATATCGAGTATG	1190 CGTTGTTAGGTAAAGCTCTGATA
	SerGluIleGluAsnArgMet	GlnAspIleGluTyrA	laLeuLeuGlyLysAlaLeuIle
	1210	1230	1250
	TTTGAAGACTCTACTGAGTAT	ATTCTGAGGCAGCTTG	CTAATTATGAGTTTAAGTGTTCT
10	PheGluAspSerThrGluTyr	IleLeuArgGlnLeuA	laAsnTyrGluPheLysCysSer
	1270	1290	1310
	CATCATAAAAACATATTCATA	GTATTTAAACACTTAA	NAGACAATGGATTACCTATAACT
	HisHisLysAsnIlePheIle'	ValPheLysHisLeuL	ysAspAsnGlyLeuProIleThr
15	1330	1350	1370
	GTAGACTCGGCTTGGGAAGAG	CTTTTGCGGCGTCGTA	TCAAAGATATGGACAAATCGTAT
	ValAspSerAlaTrpGluGlu	LeuLeuArgArgArgI	leLysAspMetAspLysSerTyr
	1390	1410	1430
20	CTCGGGTTAATGTTGCATGAT	GCTTTATCAAATGACA	AGCTTAGATCCGTTTCTCATACG
	LeuGlyLeuMetLeuHisAsp.	AlaLeuSerAsnAspL	ysLeuArgSerValSerHisThr
	1450	1470	1490
	GTTTTCCTCGATGATTTGAGC	GTGTGTAGCGCTGAAG.	AAAATTTGAGTAATTTCATTTTC
25	ValPheLeuAspAspLeuSer	ValCysSerAlaGluG	luAsnLeuSerAsnPheIlePhe
	1510	1530	1550
	CGCTCGTTTAATGAGTACAAT	GAAAATCCATTGCGTA	GATCTCCGTTTCTATTGCTTGAG
	ArgSerPheAsnGluTyrAsn	GluAsnProLeuArgA	rgSerProPheLeuLeuLeuGlu
30	1570	1590	1610
	CGTATAAAGGGAAGGCTTGAT	AGTGCTATAGCAAAGA	CTTTTTCTATTCGCAGCGCTAGA
	ArgileLysGlyArgLeuAsp	SerAlalleAlaLysT	hrPheSerIleArgSerAlaArg
	1630	1650	1670
35	GGCCGGTCTATTTATGATATA	TTCTCACAGTCAGAAA	TTGGAGTGCTGGCTCGTATAAAA
33	GlyArgSerIleTyrAspIle	PheSerGlnSerGluI	leGlyValLeuAlaArgIleLys
	1690	1710	1730
			CTTTGATGGCTTCCCAACAGGA
40	Lysargargvalalapneser	GluasnGlnasnSelP	hePheAspGlyPheProThrGly
	1750	1770	1790
	TACAAGGATATTGATGATAAA	GG AGTTATCTTAGCTA	NAAGGTAATTTCGTGATTATAGCA
	TyrLysAspIleAspAspLys	GlyVallleLeuAlaI	.ysGlyAsnPheValIleIleAla
45	1810	1830	1850
45	GCTAGACCATCTATAGGGAAA	ACAGCTTTAGCTATAG	ACATGGCGATAAATCTTGCGGTT
	AlaArgProSerIleGlyLys	ThrAlaLeuAlaIleA	spMetAlaIleAsnLeuAlaVal
	1870	1890	1910
	ACTCAACAGCGTAGAGTTGGT	TTCCTATCTCTAGAAA	TGAGCGCAGGTCAAATTGTTGAG
50	ThrGlnGlnArgArgValGly	PneLeuSerLeuGluM	letSerAlaGlyGlnIleValGlu
	1930	1950	1970
	CGGATTATTGCTAATTTAACA	AGGAATATCTGGTGAAA	<i>LAATTACAAAGAGGGGATCTCTCT</i>
	ArgileileAlaAsnLeuTh	GlyIleSerGlyGluI	LysLeuGlnArgGlyAspLeuSer

	1990	2010	2030
	AAAGAAGAATTATTCCGAGTA		
	LysGluGluLeuPheArgVal		
	Dysoluolubeurnenigval	ordoranta ory ordini	vaiki gozuset nisi ne i yi
5	2050	2070	2090
	ATCTGCAGTGATAGTCAGTAT	AAGCTTAACTTAATCGCG	AATCAGATCCGGTTGCTGAGA
	IleCysSerAspSerGlnTyr		
	11eCysSetAspbetOInty:	by suc unsube ulical u	Ashormiterrybeobeurry
	2110	2130	2150
10	AAAGAAGATCGAGTAGACGTA	ATATTTATCGATTACTTG	CAGTTGATCAACTCATCGGTT
	LysGluAspArgValAspVal		
		• •	
	2170	2190	2210
	GGAGAAAATCGTCAAAATGAA		
15	GlyGluAsnArgGlnAsnGlu	IleAlaAspIleSerArg	ThrLeuArgGlyLeuAlaSer
		2252	
	2230	2250	2270
	GAGCTAAACATTCCTATAGTT		
	GluLeuAsnIleProIleVal	CysLeuSerGlnLeuSer	ArgLysValGluAspArgAla
20	2290	2310	2330
	AATAAAGTTCCCATGCTTTCA		
	AsnLysValProMetLeuSer	AspredargaspserGry	GINITEGIUGINASPATAASP
	2350	2370	2390
25	GTGATTTTGTTTATCAATAGG	AAGGAATCGTCTTCTAAT	
25	ValileLeuPheileAsnArg		
	,		0,0000000000000000000000000000000000000
	2410	2430	2450
	AATAGACATGGATCGGTTTTC		
20	AsnArgHisGlySerValPhe	SerSerValLeuHisPhe	AspProLysIleSerLysPhe
30			
	2470	2490	2510
	TCCGCTATTAAAAAAGTATGG		'ACTTCATCAAAAGTCCTATCC
	SerAlaIleLysLysValTrp		
	ORF2 >> Metv	alAsnTyrSerAsnCysi	lisPheIle Ly sSerProIle
35 "	2530	0550	2570
	ACCTTGAAAATCAGAAGTTTG	2550	
	isLeuGluAsnGlnLysPheG	diyargargProGlyGlnS	eriterAsiteSethtorAsr
	2590	2610	2630
40	TGGCTCAAAATGGGATGGTAG		TTTCTTTCATCATTACCAT
	euAlaGlnAsnGlyMetValG	luVal IleGlyLeuAspl	PheLeuSerSerHisTvrHis
	2650	2670 .	2690
	CATTAGCAGCTATCCAAAGAT	TACTGACCGCAACGAATT	ACAAGGGGAACACAAAAGGGG
45	laLeuAlaAlaIleGlnArgi		
	_		•
	2710	2730	2750
	TTGTTTTATCCAGAGAATCA	ATAGTTTTCAATTTGAAC	GATGGATACCAAGAATCCGTT
	alValLeuSerArgGluSer#	snSerPheGlnPheGlu(GlyTrpIleProArgIleArgP
50	2770	2700	2010
		2790	. 2810
	TTACAAAACTGAATTCTTAC	SAGGCTTATGGAGTTAAG	GGTATAAAACATCCAGAAATA
	heThrLysThrGluPheLeu(SIUAlaTyrGlyValLys	ArgTyrLysThrSerArgAsnL

	2830 AGTATGAGTTTAGTGGAAAAGAA	2850 GCTGAAACTGC	2870 TTTAGAAGCCTTATACCATTTA	AGGAC
	ysTyrGluPheSerGlyLysGlu	AlaGluThrAla	Tendinalacentathisre	пСтАн
5	2890	2910	2930	
	ATCAACCGTTTTTAATAGTGGCA	ACTAGAACTCG/	ATGGACTAATGGAACACAAAT	AGTAG
	isGlnProPheLeuIleValAla	ThrArgThrAr	jTrpThrAsnGlyThrGlnIl	eValA
	2950	2970	2990	<u>.</u>
10	ACCGTTACCAAACTCTTTCTCCG.	ATCATTAGGAT'	PTACGAAGGATGGGAAGGTTT	AACTG
	spArgTyrGlnThrLeuSerPro	11e11eArg11	stAreinerAribeinerAre.	UTNTA
	3010	3030	3050	
	ACGAAGAAATATAGATATAGAC	TTAACACCTTT'	PARTICACCACCTACACGGAA	ACATA
15	spGluGluAsnIleAspIleAsp	LeuThrProPh	easnSerProProTnrArgLy:	SHISL
	3070	3090	3110	
	AAGGGTTCGTTGTAGAGCCATGT	CCTATCTTGGT	AGATCAAATAGAATCCTACTT	TGTAA
	ysGlyPheValValGluProCys	ProlleLeuVa	lAspGlnIleGluSerTyrPh	eValI
20	3130	3150	3170	
	TCAAGCCTGCAAATGTATACCAA	GAAATAAAAAT	GCGTTTCCCAAATGCATCAAA	GTATG
	leLysProAlaAsnValTyrGln	GlulleLysMe	tArgPheProAshAlaSerLy	STYTA
	3190	3210	3230	
25	CTTACACATTTATCGACTGGGTG	ATTACAGCAGC	TGCGAAAAAGAGACGAAAATT.	AACTA
	laTyrThrPheIleAspTrpVal	IleThrAlaAl	aAlaLysLysArgArgLysLe	uThrL
	3250	3270	3290	
	AGGATAATTCTTGGCCAGAAAAC	TTGTTATTAAA	CGTTAACGTTAAAAGTCTTGC	ATATA
30	YSASPASNSerTrpProGluAsn	LeuLeuLeuAs	nvalAsnvalLysSerLeuAl	aryrı
	3310	3330	3350	
	TTTTAAGGATGAATCGGTACATC	TGTACAAGGAA	CTGGAAAAAAATCGAGTTAGC	TATCG
	leLeuArgMetAsnArgTyrIle	CysThrArgAs	nTrpLysLysIleGluLeuAl	alleA
35	3370	3390	3410	
	ATAAATGTATAGAAATCGCCATT	CAGCTTGGCTG	GTTATCTAGAAGAAAACGCAT	TGAAT
	spLysCysIleGluIleAlaIle	GInLeuGIYTT	pLeuSerArgArgLySArg11	eGIUP
	3430	3450	3470	
ю	TTCTGGATTCTTCTAAACTCTCT			
	heLeuAspSerSerLysLeuSer	rysrysgiuli	eLeuTyrLeuAsnLysGluAi	rgrne
	3490	3510	3530	
	AAGAAATAACTAAGAAATCTAAA	GAACAAATGGA	ACAATTAGAACAAGAATCTAT	TTAATT
15	luGluIleThrLysLysSerLys	GluGinmetGi	uGInLeuGIuGInGIuSerii	easne
	3550	3570	3590	
	AATAGCAAGCTTGAAACTAAAAA	CCTAATTTATT	TAAAGCTCAAAATAAAAAAAGA	GTTTT
	nd			
50	3610	3630	3650	
-	AAAATGGGAAATTCTGGTTTTT	ATTTGTATAACA	CTGAAAACTGCGTCTTTGCTG	FAATA
	ORF3>> MetGlyAsnSerGlyPheTy	rLeuTyrAsn1	hrGluAsnCysValPheAlaA	AspAsn
	3670	3690	3710	
· E	ATCAAAGTTGGGCAAATGACAGA	AGCCGCTCAAGC	ACCAGCAAATAATCCTTGGGA	CAACA
5	IleLysValGlyGlnMetThrG	luProLeuLys#	\spGlnGlnIleIleLeuGly7	lhrThi

	3730	3750	3770
	TCAACACCTGTCGCAGCCA	AAATGACAGCTTCTGATG(GAATATCTTTAACAGTCTCCAAT
	SerThrProValAlaAlaL	ysMetThrAlaSerAspG]	lyIleSerLeuThrValSerAsn
5	3790	3810	3830
-			ATGCGGAAAAAGCTTACCAGCTT
	AsnSerSerThrAsnAlaS	erlleThrlleGlyLeuA:	spAlaGluLysAlaTyrGlnLeu
	3950	3870	3000
	3850	3870	3890 TTGCTGATACTATTGTTGATAGT
10	TlatauGlutustauGluk	enGlalleleukenGluti	leAlaAspThrIleValAspSer
	Trependinthapendily	spoiniteleurspoiyi	rentanspintilevalAspSet
	3910	3930	3950
			CTTCTCTAGGTTTGTTGAAAGCT
			roSerLeuGlyLeuLeuLysAla
15	-	• •	
	3970	3990	4010
			ACGGGTTATTCACTCCCAGTAAC
	PheAsnAsnPheProIleT	hrAsnLysIleGlnCysAs	snGlyLeuPheThrProSerAsn
	4030	1050	4000
20	4030	4050	4070 PCACAGTCACACCCAAAAGCTCT
	ATTGAAACTTTATTAGGAG	LAMB COLUTION COLUMN	reacagreacacecaaaagerer neThrValThrProLysSerSer
	11edidini Ledbeddiyd	TYTHIGIUITEGIYLYSPI	iernivarinterobysserser
	4090	4110	4130
	GGGAGCATGTTCTTAGTCT	CAGCAGATATTATTGCAT	CAAGAATGGAAGGCGGCGTTGTT
25			erArgMetGluGlyGlyValVal
	4150	4170	4190
	CTAGCTTTGGTACGAGAAG	GTGATTCTAAGCCCTGCG	GATTAGTTATGGATACTCATCA
	LeuAlaLeuValArgGluG	TyAspserLysProCysA	lalleSerTyrGlyTyrSerSer
30	4210	4230	4250
			4250 CTAATACAGGATTGACTCCGACA
	GlvIleProAsnLeuCvsS	erLeuArgThrSerTleTh	rAsnThrGlyLeuThrProThr
35	4270	4290	4310
55			GGTATGGGTTAATGCCCTTTCT
	ThrTyrSerLeuArgValG	lyGlyLeuGluSerGlyVa	alValTrpValAsnAlaLeuSer
	4220	4350	4334
	4330	4350	4370 Atgtatcttttttagaggtaata
40	AATOOCAATOATATTTAO	CAMIANCARATACTICIA	aigiaiciiiiitagaggiaata snValSerPheLeuGluValIle
	AsildlyAsilAspliebedo	Tylleinrasminrsera	guante chuere a dina a i i i e
	4390	4410	4430
_	CCTCAAACAAACGCTTAAA	CAATTTTTATTGGATTTT	TCTTATAGGTTTTATATTTAGAG
	ProGlnThrAsnAlaEnd		
45			
	4450	4470	4490
			AAAGAAAAGTGAGGGACGATTTT
		ORF4 >> MetGlnAsnLy	ysArgLysValArgAspAspPhe
	4510	4530	4550
50			4550 AATTAGACCTAAAAATACGAGTA
	IleLysileValivsAsnV	allusiveAenDhaDroc	luLeuAspLeuLysIleArqVal
		Ionianahtnerroo	rane anapheany attent y val
	4570	4590	4610
	AACAAGGAAAAAGTAACTT	TCTTAAATTCTCCCTTAG	AACTCTACCATAAAAGTGTCTCA
55	AsnLysGluLysValThrF	heLeuAsnSerProLeuG	luLeuTyrHisLysSerValSer

		4630	4650	4670
		CTAATTCTAGGACTGCTTCAACAAAT		でしているいましかしかしてい
5		LeuIleLeuGlyLeuLeuGlnGlnIle		
•		rediterendia pennendi unit	egransmoerbeagrybear	nerronapserrro
		4690	4710	4730
		GTTCTTGAAAAATTAGAGGATAACAG		
		ValLeuGluLysLeuGluAspAsnSe	rLeuLysLeuLysLysAlai	euilemetLeuile
10				
			4770	4790
		TTGTCTAGAAAAGACATGTTTTCCAA		CTAACGTTGGAGTT
		LeuSerArgLysAspMetPheSerLy	sAlaGluEnd	
		•	•	
			4830	4850
15		GATTTGCACACCTTAGTTTTTTGCTC		
	ORF5	>> LeuHisThrLeuValPheCysSe	rPheLysGlyGlyThrGlyI	ysThrThrLeuSer
		4870	4890	4910
		CTAAACGTGGGATGCAACTTGGCCCA	ATTTTTAGGGAAAAAAGTG1	TACTTGCTGACCTA
		LeuAsnValGlyCysAsnLeuAlaGl	nPheLeuGlyLysLysValI	euLeuAlaAspLeu
				•
20		4930	4950	4970
		GACCCGCAATCCAATTTATCTTCTGG		GTGACCAAAAAGGC
		AspProGlnSerAsnLeuSerSerGl		
			1	
		4990	5010	5030
		TTGCACGACATAGTATACACATCAAA		
25		LeuHisAspIleValTyrThrSerAs		
•		redutavahilesgriårintaerva	uwabregriage: ire.	. Jeorarur Dyebye
		5050	5070	5090
		GATAGTGTGGACCTAATTCCTGCATC		
		AspSerValAspLeuIleProAlaSe		
30		AspsetvatAspleutiertoAtase	t Lue set set Gradiusus	rigoraleanspire
-		5110	5130	5150
		CATAGAGGACCTAGTAACAACTTAAA		
		HisArgGlyProSerAsnAsnLeuLy	stentuerenwandiniki	LyskiaProPneTyr
		5130		5210
35		5170	5190	
		GACATCTGCATAATAGACACTCCACC		
		AsplleCysllelleAspThrProPr	coserrengiadiari	rAzeinvishueAgi
			5050	
		5230	5250	5270
		GCAGGAGACAAATTAATTGCTTGTTT		
40		AlaGlyAspLysLeuIleAlaCysLe	eu ThrProGluProPheSer	IleL euGly LeuGln
		5290	5310	5330
		AAGATACGTGAATTCTTAAGTTCGG		
		LysIleArgGluPheLeuSerSerVa	alGlyLysProGluGluGlu	HisIleLeuGlyIle
45				
40		5350	5370	5390
		GCTTTGTCTTTTTGGGATGATCGTA		
		AlaLeuSerPheTrpAspAspArgAs	snSerThrAsnGlnMetTyr	IleAspIleIleGlu
			-	-
		5410	5430	5450
50		TCTATTTACAAAAACAAGCTTTTTT	CAACAAAAATTCGTCGAGAT	ATTTCTCTCAGCCGT
		SerileTyrLysAsnLysLeuPheSo	erThrLysIleArqArqAsp	IleSerLeuSerArq
		• • • • • • • • • • • • • • • • • • • •	. , , , , , , , ,	•
		5470	5490	5510
		TCTCTTCTTAAAGAAGATTCTGTAG		
		SerLeuLeuLysGluAspSerValA		

	5530	5550	5570
5	ATTCTGAAGTTAACGCATGA		CATATCGAATATGAACGAGATTAC
	IleLeuLysLeuThrHisGl	ulleAlaAsnIleLeuH	isIleGluTyrGluArgAspTyr
	_		, , , , , , , , , , , , , , , , , , , ,
	5590	5610	5630
	TCTCAGAGGACAACGTGAAC	:AAACTAAAAAAAGAAGC	GGATGTCTTTTTTAAAAAAAT
10	SerGlnArgThrThrEnd		
. •	ORF6 >> ValAsn	LysLeuLysLysGluAl	.aAspValPhePheLysLysAsnG
	5.550		
	5650	5670	5690
	AAACTGCCGCTTCTCTAGAT	TTTAAGAAGACGCTTCC	CTCCATTGAACTATTCTCAGCAA
15	Ininialaalaseileuasp	PherysrysThrLeuPr	oSerIleGluLeuPheSerAlaT
	5710	5730	5750
			3/50 CATTTTTATCAGAGTCCCAAAACT
	hrLeuAsnSerGluGluSer	GlnSerLeuAsnArgLe	uPheLeuSerGluSerGlnAsnT
		oor be anophic que	di menennarolunarollusmi
20	5770	5790	5810
	ATTCGGATGAAGAATTTTAT	CAAGAAGACATCCTAGC	GGTAAAACTGCTTACTGGTCAGA
	yrSerAspGluGluPheTyr	GlnGluAspIleLeuAl	aValLysLeuLeuThrGlyGlnI
		_	•
	5830	5850	5870
25	TAAAATCCATACAGAAGCAA	CACGTACTTCTTTTAGG	AGAAAAAATCTATAATGCTAGAA
	leLysSerIleGlnLysGln	HisValLeuLeuLeuGl	yGluLysIleTyrAsnAlaArgL
	5890	5910	5030
			5930 TTCATCTTGGATAGAGTTAGTTT
	vslleLeuSerLucAcnHic	Phasarsarthrohroh	eSerSerTrpIleGluLeuValF
30	jorrebedeerbjonspins	Theserser Int Int Fit	eserserirpileGluLeuvalk
	5950	5970	5990
	TTAGAACTAAGTCTTCTGCT	TACAATGCTCTTGCATA	TTACGAGCTTTTTATAAACCTCC
	heArgThrLysSerSerAla	TyrAsnAlaLeuAlaTy	rTyrGluLeuPheIleAsnLeuP
		· •	•
35	6010	6030	6050
	CCAACCAAACTCTACAAAAA	GAGTTTCAATCGATCCC	CTATAAATCCGCATATATTTTGG
	roAsnGinThrLeuGinLys	GluPheGlnSerIlePr	oTyrLysSerAlaTyrIleLeuA
	6070	6000	
		6090	6110 GATAGGGAAAGTATGTGGAATGT
40	lablabroLysGlybsnLey	TyeThe LyeVal Acava	datagggaaagtatgtggaatgt llleGlyLysValCysGlyMetS
	rantant day sor yaspae a	Dysintrysvalkspva	TITEGIALASAGIAMEES
	6130	6150	6170
	CGAACTCATCGGCGATAAGG	GTGTTGGATCAATTTCT	TCCTTCATCTAGAAACAAAGACG
	erAsnSerSerAlaIleArg	ValLeuAspGlnPheLe	uProSerSerArgAsnLysAspV
45	•	-	3 ·
	6190	6210	6230
	TTAGAGAAACGATAGATAAG	TCTGATTCAGAGAAGAA	TCGCCAATTATCTGATTTCTTAA
	alArgGluThrIleAspLys	SerAspSerGluLysAs	nArgGlnLeuSerAspPheLeuI
50	6250	6270	6290
	INGAGATACTTCGCATCATG	TGTTCCGGAGTTTCTTT	GTCCTCCTATAACGAAAATCTTC
	reginitenenwiditewet	claseicilaraizeire	uSerSerTyrAsnGluAsnLeuL
	6310	6330	6350
			6550 ATCCTCGTCAGCTCATATATATA
55	euGlnGlnLeuPheGluLeu		

5	6370	6390	6410
•	ATATCTATTATATATATATAT	\TTTAGGGATTTGATTTC!	ACGAGAGAGATTTGCAACTCTTG
	6430	6450	6470
	GTGGTAGACTTTGCAACTC'	ltggtggtagactttgcai	ACTCTTGGTGGTAGACTTTGCAA
10	6490	6510	6530
	CTCTTGGTGGTAGACTTGG'	ICATAATGGACTTTTGTT	AAAAATTTATTAAAATCTTAGA
	6550	6570	6590
			PCGATGGCTTTCCATAAAAGTAG SermetalaPhoHisLysSerAr
15			seruerwratueur ph i b oerwr
	6610	6630	6650
	gLeuPheLeuThrPheGly	JACGCGTCGGAAATTTGG AspAlaSerGluIleTrpI	TTATCTACTTTATCTTATCTAAC LeuSerThrLeuSerTyrLeuTh
20	6670	6690	6710
			ICTTTAGAGATTCTGGATTTATC
	rArgLysAsnTyrAlaSer	31y11eAsnPheLeuVal:	SerLeuGluIleLeuAspLeuSe
	6730	6750	6770
25			GAATCTTTGTTTAAAATCAAGTC GluserLeuPheLysIleLysSe
	6790	6810	6830
			TCTAAACAGGCTAGAGCGGCATG
30	redaspoareneasigry	rAenathatsetetmyta:	SerLysGlnAlaArgAlaAlaCy
	6850	6870	6890
			AAGGGATATATTAAACCCGCTAT LysGlyTyrileLysProAlail
35	6910	6930	6950
			ATCCGAGACAAAATCAAAACAGA IleArgAspLysIleLysThrGl
	6970	6990	7010
40			GCGCTCCGGATAGTGAATTATAG
	userileserLysGinGlu	TrpTnrvalPnePneGlu	AlaLeuArgIleValAsnTyrAr
	7030	7050	7070
	AGACTATTTAATCGGTAAA	TTGATTGTACAAGGGATC	CGTAAGTTAGACGAAATTTTGTC
45	gAspTyrLeulleGlyLys	renilevalGinGiyile	ArgLysLeuAspGluIleLeuSe
	7090	7110	7130
			ATTTCCTTTCGCATTAAAAAAA
	rLeuArgThrAspAspLeu	PhePheAlaSerAsnGln	IleSerPheArgIleLysLysAi
50	7150	7170	7190
	ACAGAATAAAGAAACCAAA gGlnAsnLysGluThrLys	ATTCTAATCACATTTCCT IleLeuIleThrPhePro	ATCAGCTTAATGGAAGAGTTGC IleSerLeuMetGluGluLeuG
	7210	7230	7250
<i></i>	AAAATACACTTGTGGGAGA	AATGGGAGAGTATTTGTT	TCTAAAATAGGGATTCCTGTAA
55	nLysTyrThrCysGlyArg	AsnGlyArgValPheVal	SerLysIleGlyIleProValT

7270 7290 7310
AACAAGTCAGGTTGCGCATAATTTTAGGCTTGCAGAGTTCCATAGTGCTATGAAAATAAA
rThrSerGlnValAlaHisAsnPheArgLeuAlaGluPheHisSerAlaMetLysIleLy

7330 7350 7370
AATTACTCCCAGAGTACTTCGTGCAAGCGCTTTGATTCATTTAAAGCAAATAGGATTAAA
sileThrProArgValLeuArgAlaSerAlaLeuIleHisLeuLysGlnIleGlyLeuLy

7450 7470 7490
TTCTGGGGAAGAGTAATTCCTCTAGTACAAACACCCACAATATTGTGATATAATTAAAA
sSerGlyGluGluValIleProLeuValGlnThrProThrIleLeuEnd

²⁰ TT

10

15

- pGO plasmid constituted by the pUC8 recombinant plasmid containing an insert corresponding to the nucleotidic sequence as per claim 1, cloned in the Barn H1 site.
 - 3. Escherichia coli transformed with the plasmid according to claim 2 and deposited as ATCC 68314.
- 4. ORF1D gene characterized by the nucleotidic sequence comprised between 1129 and 2481 in the nucleotidic sequence according to claim 1.
 - ORF2D gene characterized by the nucleotidic sequence comprised between 2480 and 3539 in the nucleotidic sequence according to claim 1.
- 35 6. ORF3D gene characterized by the nucleotidic sequence comprised between 3604 and 4395 in the nucleotidic sequence according to claim 1.
 - ORF4D gene characterized by the nucleotidic sequence comprised between 4468 and 4773 in the nucleotidic sequence according to claim 1.
 - 8. ORF5D gene characterized by the nucleotidic sequence comprised between 4804 and 5595 in the nucleotidic sequence according to claim 1.
- 9. ORF6D gene characterized by the nucleotidic sequence comprised between 5595 and 6335 in the nucleotidic sequence according to claim 1.
 - 10. ORF7D gene characterized by the nucleotidic sequence comprised between 6560 and 7486 in the nucleotidic sequence according to claim 1.
- 50 11. ORF8D gene characterized by the nucleotidic sequence complemental to the one comprised between 41 and 1030 in the nucleotidic sequence according to claim 1.
 - 12. Protein expressed by the gene according to claim 4 and characterized by the following aminoacid sequence:

55

pgp1:

5

15

20

25

MetLysThrArgSerGluIleGluAsnArgMetGlnAspIleGluTyrAlaLeuLeuGly LysAlaLeuIlePheGluAspSerThrGluTyrIleLeuArgGlnLeuAlaAsnTyrGlu PheLysCysSerHisHisLysAsnIlePheIleValPheLysHisLeuLysAspAsnGly 10 ${\tt LeuProIleThrValAspSerAlaTrpGluGluLeuLeuArgArgArgIleLysAspMet}$ AspLysSerTyrLeuGlyLeuMetLeuHisAspAlaLeuSerAsnAspLysLeuArgSer ValSerHisThrValPheLeuAspAspLeuSerValCysSerAlaGluGluAsnLeuSer AsnPhellePheArgSerPheAsnGluTyrAsnGluAsnProLeuArgArgSerProPhe LeuLeuLeuGluGlyArgSerIleTyrAspIlePheSerGlnSerGluIleGlyValLeu AlaArgIleLysLysArgArgValAlaPheSerGluAsnGlnAsnSerPhePheAspGly PheProThrGlyTyrLysAspIleAspAspLysGlyValIleLeuAlaLysGlyAsnPhe ValllelleAlaAlaArgProSerIleGlyLysThrAlaLeuAlaIleAspMetAlaIle AsnLeuAlaValThrGlnGlnArgArgValGlyPheLeuSerLeuGluMetSerAlaGly ${\tt GlnIleValGluArgIleIleAlaAsnLeuThrGlyIleSerGlyGluLysLeuGlnArg}$

30

35

40

45

GlyAspLeuSerLysGluGluLeuPheArgValGluGluAlaGlyGluThrValArgGlu ${\tt SerHisPheTyrIleCysSerAspSerGInTyrLysLeuAsnLeuIleAlaAsnGlnIle}$ ArgLeuLeuArgLysGluAspArgValAspValIlePheIleAspTyrLeuGlnLeuIle AsnSerSerValGlyGluAsnArgGlnAsnGluIleAlaAspIleSerArgThrLeuArg GlyLeuAlaSerGluLeuAsnIleProIleValCysLeuSerGlnLeuSerArgLysVal GluAspArgAlaAsnLysValProMetLeuSerAspLeuArgAspSerGlyGlnIleGlu **GlnAspAlaAspValIleLeuPheIleAsnArgLysGluSerSerSerAsnCysGluIle** ThrValGlyLysAsnArgHisGlySerValPheSerSerValLeuHisPheAspProLys IleSerLysPheSerAlaIleLysLysValTrpEnd

50

or parts of it.

13. Protein expressed by the gene according to claim 5 and characterized by the following aminoacid sequence:

pgp2:

PheGlyArgArgProGlyGlnSerIleLysIleSerProIleHisLeuGluAsnGlnLys
PheGlyArgArgProGlyGlnSerIleLysIleSerProLysLeuAlaGlnAsnGlyMet
ValGluValIleGlyLeuAspPheLeuSerSerHisTyrHisAlaLeuAlaAlaIleGln
ArgLeuLeuThrAlaThrAsnTyrLysGlyAsnThrLysGlyValValLeuSerArgGlu
SerAsnSerPheGlnPheGluGlyTrpIleProArgIleArgPheThrLysThrGluPhe
LeuGluAlaTyrGlyValLysArgTyrLysThrSerArgAsnLysTyrGluPheSerGly
LysGluAlaGluThrAlaLeuGluAlaLeuTyrHisLeuGlyHisGlnProPheLeuIle
ValAlaThrArgThrArgTrpThrAsnGlyThrGlnIleValAspArgTyrGlnThrLeu
SerProIleIleArgIleTyrGluGlyTrpGluGlyLeuThrAspGluGluAsnIleAsp
IleAspLeuThrProPheAsnSerProProThrArgLysHisLysGlyPheValValGlu
ProCysProIleLeuValAspGlnIleGluSerTyrPheValIleLysProAlaAsnVal
TyrGlnGluIleLysMetArgPheProAsnAlaSerLysTyrAlaTyrThrPheIleAsp

GluAsnLeuLeuLeuAsnValAsnValLysSerLeuAlaTyrIleLeuArgMetAsnArg

TyrIleCysThrArgAsnTrpLysLysIleGluLeuAlaIleAspLysCysIleGluIle

AlaIleGlnLeuGlyTrpLeuSerArgArgLysArgIleGluPheLeuAspSerSerLys

LeuSerLysLysGluIleLeuTyrLeuAsnLysGluArgPheGluGluIleThrLysLys

SerLysGluGlnMetGluGlnLeuGluGlnGluSerIleAsnEnd

or parts of it.

14. Protein expressed by the gene according to claim 6 and characterized by the following aminoacid sequence:

pgp3:

MetGlyAsnSerGlyPheTyrLeuTyrAsnThrGluAsnCysValPheAlaAspAsnIle 5 LysValGlyGlnMetThrGluProLeuLysAspGlnGlnIleIleLeuGlyThrThrSer ThrProValAlaAlaLysMetThrAlaSerAspGlyIleSerLeuThrValSerAsnAsn SerSerThrAsnAlaSerIleThrIleGlyLeuAspAlaGluLysAlaTyrGlnLeuIle 10 LeuGluLysLeuGlyAspGlnIleLeuAspGlyIleAlaAspThrIleValAspSerThr ValGlnAspIleLeuAspLysIleLysThrAspProSerLeuGlyLeuLeuLysAlaPhe AsnAsnPheProIleThrAsnLysIleGlnCysAsnGlyLeuPheThrProSerAsnIle 15 GluThrLeuLeuGlyGlyThrGluIleGlyLysPheThrValThrProLysSerSerGly SerMetPheLeuValSerAlaAspIleIleAlaSerArgMetGluGlyGlyValValLeu 20 AlaLeuValArgGluGlyAspSerLysProCysAlaIleSerTyrGlyTyrSerSerGly IleProAsnLeuCysSerLeuArgThrSerIleThrAsnThrGlyLeuThrProThrThr TyrSerLeuArgValGlyGlyLeuGluSerGlyValValTrpValAsnAlaLeuSerAsn 25 GlyAsnAspIleLeuGlyIleThrAsnThrSerAsnValSerPheLeuGluValIlePro GlnThrAsnAlaEnd

30

or parts of it.

sequence:

35

40

45

pgp4:

MetGlnAsnLysArgLysValArgAspAspPheIleLysIleValLysAspValLysLys
AspPheProGluLeuAspLeuLysIleArgValAsnLysGluLysValThrPheLeuAsn
SerProLeuGluLeuTyrHisLysSerValSerLeuIleLeuGlyLeuLeuGlnGlnIle
GluAsnSerLeuGlyLeuPheProAspSerProValLeuGluLysLeuGluAspAsnSer
LeuLysLeuLysLysAlaLeuIleMetLeuIleLeuSerArgLysAspMetPheSerLys
AlaGluEnd

15. Protein expressed by the gene according to claim 7 and characterized by the following aminoacid

50

or parts of it.

55

16. Protein expressed by the gene according to claim 8 and characterized by the following aminoacid sequence:

pgp5:

LeuHisThrLeuValPheCysSerPheLysGlyGlyThrGlyLysThrThrLeuSerLeu AsnValGlyCysAsnLeuAlaGlnPheLeuGlyLysLysValLeuLeuAlaAspLeuAsp ProGlnSerAsnLeuSerSerGlyLeuGlyAlaSerValArgSerAspGlnLysGlyLeu HisAspIleValTyrThrSerAsnAspLeuLysSerIleIleCysGluThrLysLysAsp 10 SerValAspLeuIleProAlaSerPheSerSerGluGlnPheArgGluLeuAspIleHis ArgGlyProSerAsnAsnLeuLysLeuPheLeuAsnGluTyrCysAlaProPheTyrAsp 15 IleCysIleIleAspThrProProSerLeuGlyGlyLeuThrLysGluAlaPheValAla **GlyAspLysLeuIleAlaCysLeuThrProGluProPheSerIleLeuGlyLeuGlnLys** IleArgGluPheLeuSerSerValGlyLysProGluGluGluHisIleLeuGlyIleAla 20 LeuSerPheTrpAspAspArgAsnSerThrAsnGlnMetTyrIleAspIleIleGluSerIleTyrLysAsnLysLeuPheSerThrLysIleArgArgAspIleSerLeuSerArgSer LeuLeuLysGluAspSerValAlaAsnValTyrProAsnSerArgAlaAlaGluAspIle LeuLysLeuThrHisGluIleAlaAsnIleLeuHisIleGluTyrGluArgAspTyrSer GlnArgThrThrEnd 30

or parts of it.

40

50

55

17. Protein expressed by the gene according to claim 9 and characterized by the following aminoacid sequence:

pgp6:

 ${\tt ValAsnLysLeuLysLysGluAlaAspValPhePheLysLysAsnGlnThrAlaAlaSer}$ ${\tt LeuAspPheLysLysThrLeuProSerIleGluLeuPheSerAlaThrLeuAsnSerGlu}$ 5 ${\tt GluSerGlnSerLeuAspArgLeuPheLeuSerGluSerGlnAsnTyrSerAspGluGlu}$ ${\tt PheTyrGlnGluAspIleLeuAlaValLysLeuLeuThrGlyGlnIleLysSerIleGln}$ 10 LysGlnHisValLeuLeuGlyGluLysIleTyrAsnAlaArgLysIleLeuSerLys **AspHisPheSerSerThrThrPheSerSerTrpIleGluLeuValPheArgThrLysSer** SerAlaTyrAsnAlaLeuAlaTyrTyrGluLeuPhelleAsnLeuProAsnGlnThrLeu 15 GlnLysGluPheGlnSerIleProTyrLysSerAlaTyrIleLeuAlaAlaArgLysGly ${\tt AspLeuLysThrLysValAspValIleGlyLysValCysGlyMetSerAsnSerSerAla}$ 20 ${\tt IleArgValLeuAspGlnPheLeuProSerSerArgAsnLysAspValArgGluThrIle}$ ${\tt AspLysSerAspSerGluLysAsnArgGlnLeuSerAspPheLeuIleGluIleLeuArg}$ ${\tt IleMetCysSerGlyValSerLeuSerSerTyrAsnGluAsnLeuLeuGlnGlnLeuPhe}$ 25 GluLeuPheLysGlnLysSerEnd

or parts of it.

30

35

45

18. Protein expressed by the gene according to claim 10 and characterized by the following aminoacid sequence:

pgp7 :

LeuValLysLysMetGlySerMetAlaPheHisLysSerArgLeuPheLeuThrPheGly AspAlaSerGluIleTrpLeuSerThrLeuSerTyrLeuThrArgLysAsnTyrAlaSer

40 GlyIleAsnPheLeuValSerLeuGluIleLeuAspLeuSerGluThrLeuIleLysAla
IleSerLeuAspHisSerGluSerLeuPheLysIleLysSerLeuAspValPheAsnGly

55

LysValValSerGluAlaSerLysGlnAlaArgAlaAlaCysTyrIleSerPheThrLys
PheLeuTyrArgLeuThrLysGlyTyrIleLysProAlaIleProLeuLysAspPheGly
AsnThrThrPhePheLysIleArgAspLysIleLysThrGluSerIleSerLysGlnGlu
TrpThrValPhePheGluAlaLeuArgIleValAsnTyrArgAspTyrLeuIleGlyLys
LeuIleValGlnGlyIleArgLysLeuAspGluIleLeuSerLeuArgThrAspAspLeu
PhePheAlaSerAsnGlnIleSerPheArgIleLysLysArgGlnAsnLysGluThrLys
IleLeuIleThrPheProIleSerLeuMetGluGluLeuGlnLysTyrThrCysGlyArg
AsnGlyArgValPheValSerLysIleGlyIleProValThrThrSerGlnValAlaHis
AsnPheArgLeuAlaGluPheHisSerAlaMetLysIleLysIleThrProArgValLeu
ArgAlaSerAlaLeuIleHisLeuLysGlnIleGlyLeuLysAspGluGluIleMetArg
IleSerCysLeuSerSerArgGlnSerValCysSerTyrCysSerGlyGluGluValIle
ProLeuValGlnThrProThrIleLeuEnd

25 or parts of it.

19. Protein expressed by the gene according to claim 11 and characterized by the following aminoacid sequence:

pgp8:

MetGlyLysGlyIleLeuSerLeuGlnGlnGluMetSerLeuGluTyrSerGluLysSer
TyrGlnGluValLeuLysIleArgGlnGluSerTyrTrpLysArgMetLysSerPheSer
LeuPheGluValIleMetHisTrpThrAlaSerLeuAsnLysHisThrCysArgSerTyr
ArgGlySerPheLeuSerLeuGluLysIleGlyLeuLeuSerLeuAspMetAsnLeuGln
GluPheSerLeuLeuAsnHisAsnLeuIleLeuAspAlaIleLysLysValSerSerAla
LysThrSerTrpThrGluGlyThrLysGlnValArgAlaAlaSerTyrIleSerLeuThr
ArgPheLeuAsnArgMetThrGlnGlyIleValAlaIleAlaGlnProSerLysGlnGlu
AsnSerArgThrPhePheLysThrArgGluIleValLysThrAspAlaMetAsnSerLeu
GlnThrAlaSerPheLeuLysGluLeuLysLysIleAsnAlaArgAspTrpLeuIleAla
GlnThrMetLeuGlnGlyGlyLysArgSerSerGluValLeuSerLeuGluIleSerGln
IleCysPheGlnGlnAlaThrIleSerPheSerGlnLeuLysAsnArgGlnThrGluLys
ArgIleIleIleThrTyrProGlnLysPheMetHisPheLeuGlnGluTyrIleGlyGln

ArgArgGlyPheValPheValThrArgSerGlyLysMetValGlyLeuArgGlnIleAla
ArgThrPheSerGlnAlaGlyLeuGlnAlaAlaIleProPheLysIleThrProHisVal
LeuArgAlaThrAlaValThrGluTyrLysArgLeuGlyCysSerAspSerAspIleMet
LysValThrGlyHisAlaThrAlaLysMetIlePheAlaTyrAspLysSerSerArgGlu
AspAsnAlaSerLysLysMetAlaLeuIleEnd

or parts of it.

10

15

20

25

- 20. Recombinant expression vectors characterized by containing the genes according to claims 4-11.
- 21. Expression vector according to claim 20 in which the vector pertains to the pEX34 family, the cloned insert is a gene according to claims 4-11, the host cell is E.coli K12ΔH1Δtrp.
- 22. pO3/GO/MC1 plasmid, constituted by the recombinant expression vector pEX34 and a ORF3D insert.
- 23. Escherichia coli transformed with the recombinant expression vector according to claim 22 and deposited as ATCC 68315.
- 24. Process for preparing the immunogenic protein according to claims 12-19 in which:
 - a) an ORF is isolated according to claims 4-11
 - b) said ORF is cloned in an expression vector and the thus obtained recombinant vector is isolated
 - c) bacterial cells are transformed with the aid of a recombinant vector of stage (b)
 - d) the bacterial cells transformed as in (c) are cultivated in a suitable medium
 - e) the thus obtained protein is isolated and purified from the cell lysate.

30

- 25. Process according to claim 24 in which the vector as per stage (b) is pEX34.
- 26. Process according to claim 25 in which the ORF as per stage (a) is ORF3D.
- 27. Process according to claim 26 in which the cells as per stage (d) are the ones deposited as ATCC 68315 and the protein product is a recombinant protein (MS2-pgp3) constituted by a terminal portion generated by the vector and by the portion of the pgp3D protein.
- 28. Process according to claim 27 in which the cell lysate obtained from strain ATCC 68315 is partially purified by dialysis against a phosphate buffer consisting of 0.4% KCI, 0.4% KH₂PO₄, 16% NaCI, 2.5% NaH₂PO₄ at 4°C for about 15 hours, the thus obtained precipitate is discarded and the protein solution is utilized both as such as an antigen in diagnostic tests and further purified.
- 29. Recombinant MS2-pgp3D protein resulting from the process according to claim 26 and represented by the aminoacid sequence:

50

-106
MetSerLysThrThrLysLysPheAsnSerLeuCysIleAspLeuProArgAspLeuSer
LeuGluIleTyrGlnSerIleAlaSerValAlaThrGlySerGlyAspProHisSerAsp
AspPheThrAlaIleAlaTyrLeuArgAspGluLeuLeuThrLysHisProThrLeuGly
SerGlyAsnAspGluAlaThrArgArgThrLeuAlaIleAlaLysLeuArgGluAlaAsn
GlyAspArgGlyGlnIleAsnArgGluGlyPheLeuHisAspLysSerLeuSerTrpAsp
+1
IleArgAlaThrGlySerMetGlyAsnSerGlyPheTyrLeuTyrAsnThrGluAsnCys
ValPheAlaAspAsnIleLysValGlyGlnMetThrGluProLeuLysAspGlnGlnIle
IleLeuGlyThrThrSerThrProValAlaAlaLysMetThrAlaSerAspGlyIleSer
LeuThrValSerAsnAsnSerSerThrAsnAlaSerIleThrIleGlyLeuAspAlaGlu
LysAlaTyrGlnLeuIleLeuGluLysLeuGlyAspGlnIleLeuAspGlyIleAlaAsp

ThrileValAspSerThrValGlnAspIleLeuAspLysIleLysThrAspProSerLeu
GlyLeuLeuLysAlaPheAsnAsnPheProIleThrAsnLysIleGlnCysAsnGlyLeu
PheThrProSerAsnIleGluThrLeuLeuGlyGlyThrGluIleGlyLysPheThrVal
ThrProLysSerSerGlySerMetPheLeuValSerAlaAspIleIleAlaSerArgMet
GluGlyGlyValValLeuAlaLeuValArgGluGlyAspSerLysProCysAlaIleSer
TyrGlyTyrSerSerGlyIleProAsnLeuCysSerLeuArgThrSerIleThrAsnThr
GlyLeuThrProThrThrTyrSerLeuArgValGlyGlyLeuGluSerGlyValValTrp
ValAsnAlaLeuSerAsnGlyAsnAspIleLeuGlyIleThrAsnThrSerAsnValSer
PheLeuGluValIleProGlnThrAsnAlaEnd

or parts thereof.

- 30. Vaccine against infections caused by Chlamydia trachomatis containing an immunologically effective amount of one of the proteins according to claims 12-19 and 29 and a pharmaceutically acceptable diluent.
 - 31. Vaccine according to claim 30 in which the protein is the one according to claim 14.
- 50 32. Vaccine according to claim 30 in which the protein is MS2-pgp3D2.
 - 33. Kit for immunological RIA or ELISA assays in which the antigen utilized in the search for specific antibodies to Chlamydia trachomatis is the protein according to claim 29.

55

5

10

15

20

25

30

FIG. 1A (1)

10	30	50
ATATTCATATTCTGTTGC	CAGAAAAACACCTTTAGGC1	ATATTAGAGCCATCTTCTTTG
70	90 KTAAAAAAAAAAAAAAAAA	110 ATCATCTTTGCGGTTGCGTGTC
130 CTGTGACCTTCATTATGT	150 CGGAGTCTGAGCACCCTAGGC	170 GTTTGTACTCCGTCACAGCGG
	210	230
190 TTGCTCGAAGCACGTGCG		6CAGCTTGTAGTCCTGCTTGAG
250	270	290
	GCCTTAACCCCACCATTTTTC	CGGAGCGAGTTACGAAGACAA
310	330	350
AACCTCTTCGTTGACCGA	TGTACTCTTGTAGAAAGTGCA	TAAACTTCTGAGGATAAGTTA
370	390	410
TAATAATCCTCTTTTCTG		GAGAAAGAAATGGTAGCTTGTT
430 GGAAACAAATCTGACTAA	450 TCTCCAAGCTTAAGACTTCAG	470 BAGGAGCGTTTACCTCCTTGGA
	510	530
490 GCATTGTCTGGGCGATCA		TTTTTAGCTCTTTTAGGNAGG
550	570	590
	TCATCGCATCCGTTTTTACTA	TTTCCCTGGTTTTAAAAATG
610	630	650
TTCGACTATTTTCTTGTT	TAGAAGGTTGCGCTATAGCG	ACTATTCCTTGAGTCATCCTGT
670	690	710 ACTTGTTTAGTACCTTCGGTCC
730 AAGAAGTCTTGGCAGAGG	750 Saaacttttttaatcgcatctæ	770 AGGATTAGATTATGATTTAAAA
790	810	. 830
		CCAATCTTTTCTAAAGACAAAA
850	870	890
AAGATCCTCGATATGATC	CTACAAGTATGTTGTTGAGT	GATGCGGTCCAATGCATAATAA
910	930	950
CTTCGAATAAGGAGAAGG	CTTTTCATGCGTTTCCAATAG	GATTCTTGGCGAATTTTTAAAA
970 CTTCCTCATAAGACTTTT	990 "דד גר א בר א	1010 CTTGCTGCAAAGATAAATCC
1030 CTTTACCCATGAAATCC	1050 CTCGTGATATAACCTATCCGT	1070 Aaaatgtcctgattagtgaaat
1090	1110	1130
		PTATATAAACATGAAAAC TCGT
		OPPI II Mattuethrare

FIG. 1A (2)

1150	1170	1190
TCCGAAATAGAAAATCGC	ATGCAAGATATCGAGTATGC	GTTGTTAGGTAAAGCTCTGATA
SerGluIleGluAsnArgi	MetGlnAspIleGluTyrAl.	aLeuLeuGlyLysAlaLeuIle
-	-	• •
1210	1230	1250 .
		TAATTATGAGTTTAAGTGTTCT
PheGluAspSerThrGlu'	TyrlleLeuArgGlnLeuAl	aAsnTyrGluPheLysCysSer
•	-	- ,
1270	1290	1310
		AGACAATGGATTACCTATAACT
HisHisLysAsnIlePhe:	IleValPheLysHisLeuLy:	sAspAsnGlyLeuProIleThr
		<u>-</u>
1330	1350	1370
		CAAAGATATGGACAAATCGTAT
ValAspSerAlaTrpGlu	GluLeuLeuArgArgArgIl	eLysAspMetAspLysSerTyr
1390	1410	1430
		GCTTAGATCCGTTTCTCATACG
LeuGlyLeuMetLeuHis	AspAlaLeuSerAsnAspLy:	sLeuArgSerValSerHisThr
1450	1470	1490
		AAATTTGAGTAATTTCATTTTC
ValPheLeuAspAspLeu:	SerValCysSerAlaGluGl	uAsnLeuSerAsnPheIlePhe
1510	1530	1550
CGCTCGTTTAATGAGTAC	AATGAAAATCCATTGCGTAG	ATCTCCGTTTCTATTGCTTGAG
ArgSerPheAsnGluTyr	AsnGluAsnProLeuArgAr	gSerProPheLeuLeuLeuGlu
1 - 7 - 7	1 600	4.44
1570	1590	1610
CGTATAAAGGGAAGGCTT	GATAGTGCTATAGCAAAGAC	TTTTTCTATTCGCAGCGCTAGA
CGTATAAAGGGAAGGCTT	GATAGTGCTATAGCAAAGAC	
CGTATAAAGGGAAGGCTTG ArgileLysGlyArgLeu	GATAGTGCTATAGCAAAGAC AspSerAlaIleAlaLysTh	TTTTTCTATTCGCAGCGCTAGA rPheSerIleArgSerAlaArg
CGTATAAAGGGAAGGCTTGArgIleLysGlyArgLeu	GATAGTGCTATAGCAAAGAC AspSerAlaIleAlaLysTh	TTTTTCTATTCGCAGCGCTAGA rPheSerIleArgSerAlaArg 1670
CGTATAAAGGGAAGGCTT ArgileLysGlyArgLeu 1630 GGCCGGTCTATTTATGAT	GATAGTGCTATAGCAAAGAC AspSerAlaIleAlaLysTh 1650 ATATTCTCACAGTCAGAAAT	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA
CGTATAAAGGGAAGGCTT ArgileLysGlyArgLeu 1630 GGCCGGTCTATTTATGAT	GATAGTGCTATAGCAAAGAC AspSerAlaIleAlaLysTh 1650 ATATTCTCACAGTCAGAAAT	TTTTTCTATTCGCAGCGCTAGA rPheSerIleArgSerAlaArg 1670
CGTATAAAGGGAAGGCTTGATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATGGLYATGSETILETYTASP	GATAGTGCTATAGCAAAGAC AspSerAlaIleAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluIl	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys
CGTATAAAGGGAAGGCTTC ArgileLysGlyArgLeux 1630 GGCCGGTCTATTTATGAT. GlyArgSerileTyrAsp 1690	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATGCLYATGATGATGATGATGATGATGATGATGATGATGATGATGA	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil 1710 CTGAGAATCAAAATTCTTTC	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATGCLYATGATGATGATGATGATGATGATGATGATGATGATGATGA	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil 1710 CTGAGAATCAAAATTCTTTC	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATGCLYATGATGATGATGATGATGATGATGATGATGATGATGATGA	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil 1710 CTGAGAATCAAAATTCTTTC	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCGLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCT LYSATGATGValalaPhe 1750	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil 1710 CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPho	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCATCATCATCATCATCATCATCATCATCATCATCATCA	GATAGTGCTATAGCAAAGACAASACACAASACACAASACAAAAAAAAAA	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCATCATCATCATCATCATCATCATCATCATCATCATCA	GATAGTGCTATAGCAAAGACAASACACAASACACAASACAAAAAAAAAA	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCGLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCTLYSATGATGATGATGTTCTTTSOTTACAAGGATATTGATGATTTTLYSASPILEASPASP 1810	GATAGTGCTATAGCAAAGAC ASpSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheVallleIleAla 1850
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCGLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCTLYSATGATGATGATGTTCTTTSOTTACAAGGATATTGATGATTTTLYSASPILEASPASP 1810	GATAGTGCTATAGCAAAGAC ASpSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830	TTTTTCTATTCGCAGCGCTAGA rPheSerIleArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheValIleIleAla
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUS 1630 GGCCGGTCTATTTATGATCGLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCTLYSATGATGATGATGTTCTTTSOTACAAGGATATTGATGATGTTYTLYSASPILEASPASP 1810 GCTAGACCATCTATAGGG	GATAGTGCTATAGCAAAGACAASACACAASACACAASACACAASACACAAAATCACAAAAATCTTTCCACAGAAAATCTTTCCACAGAAAATCTTTCCACAAAAAACAGCATAACAAAACAGCTTAAGCTAAAAAAAA	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheVallleIleAla 1850
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCGLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCTLYSATGATGATGATGTTGATGATGATGATGATGATGATGATGA	GATAGTGCTATAGCAAAGACAASACACAASACACAASACACAAAATCCTCACAGTCAGAAAATCTTTCCTGAGAAATCAAAAATCTTTCCTGAGAAATCAAAAATCTTTCCACAGAAATCATAGCAAAACAGCAGTTATCTAGCTAAAACAGCTTTAGCTAAAAACAGCTTTAGCTAAAAAACAGCTTTAGCTAAAAAAACAGCTTTAGCTAAAAAAACAGCTTTAGCTAAAAAAACAGCTTTAGCTAAAAAAAA	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheValileIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGATCGLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCTLYSATGATGATGATGTTCTTTSOTTACAAGGATATTGATGATGTTYTLYSASPILEASPASP 1810 GCTAGACCATCTATAGGGALAATGPTOSETILEGLY 1870	GATAGTGCTATAGCAAAGACAASACACASPSerAlaileAlaLysThe 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluile 1710 CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830 AAAACAGCTTTAGCTATAGAA LysThrAlaLeuAlaileAs	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheValileIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal 1910
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGAT. GLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCT LYSATGATGVALALAPhe 1750 TACAAGGATATTGATGAT. TYTLYSASPILEASPASP 1810 GCTAGACCATCTATAGGGALAATGPTOSETILEGLY 1870 ACTCAACAGCGTAGAGTT	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830 AAAACAGCTTTAGCTATAGAA LysThrAlaLeuAlaileAs	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheVallleIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal 1910 GAGCGCAGGTCAAATTGTTGAG
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGAT. GLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCT LYSATGATGVALALAPhe 1750 TACAAGGATATTGATGAT. TYTLYSASPILEASPASP 1810 GCTAGACCATCTATAGGGALAATGPTOSETILEGLY 1870 ACTCAACAGCGTAGAGTT	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830 AAAACAGCTTTAGCTATAGAA LysThrAlaLeuAlaileAs	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheValileIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal 1910
1630 GGCCGGTCTATTTATGAT. GlyArgSerIleTyrAsp 1690 AAAAGACGAGTAGCGTTCT LySArgArgValAlaPhe 1750 TACAAGGATATTGATGAT. TyrLySAspIleAspAsp 1810 GCTAGACCATCTATAGGG AlaArgProSerIleGly 1870 ACTCAACAGCGTAGAGTT ThrGlnGlnArgArgVal	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluile CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830 AAAACAGCTTTAGCTATAGAA LysThrAlaLeuAlaileAs LysThrAlaLeuAlaileAs GGTTTCCTATCTCTAGAAAT GlyPheLeuSerLeuGluMe	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheVallleIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal 1910 GAGCGCAGGTCAAATTGTTGAG
CGTATAAAGGGAAGGCTTA ArgileLysGlyArgLeus 1630 GGCCGGTCTATTTATGATA GlyArgSerileTyrAsp 1690 AAAAGACGAGTAGCGTTCT LysArgArgValAlaPhe 1750 TACAAGGATATTGATGATA TyrLysAspileAspAsp 1810 GCTAGACCATCTATAGGG AlaArgProSerileGly 1870 ACTCAACAGCGTAGAGTT ThrGlnGlnArgArgVal	GATAGTGCTATAGCAAAGAC AspSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830 AAAACAGCTTTAGCTATAGAA LysThrAlaLeuAlaileAs 1890 GGTTTCCTATCTCTAGAAAT GlyPheLeuSerLeuGluMe	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheValileIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal 1910 GAGCGCAGGTCAAATTGTTGAG tSerAlaGlyGlnIleValGlu 1970
CGTATAAAGGGAAGGCTTCATGILELYSGLYATGLEUG 1630 GGCCGGTCTATTTATGAT. GLYATGSETILETYTASP 1690 AAAAGACGAGTAGCGTTCT LYSATGATGVALALAPhe 1750 TACAAGGATATTGATGAT. TYTLYSASPILEASPASP 1810 GCTAGACCATCTATAGGGALAATGPTOSETILEGLY 1870 ACTCAACAGCGTAGAGTT ThrGLnGLnATGATGYVAL	GATAGTGCTATAGCAAAGAC ASpSerAlaileAlaLysTh 1650 ATATTCTCACAGTCAGAAAT IlePheSerGlnSerGluil CTGAGAATCAAAATTCTTTC SerGluAsnGlnAsnSerPhe 1770 AAAGGAGTTATCTTAGCTAA LysGlyValileLeuAlaLy 1830 AAAACAGCTTTAGCTATAGAA LysThrAlaLeuAlaileAs LysThrAlaLeuAlaileAs GGTTTCCTATCTCTAGAAAT GlyPheLeuSerLeuGluMe 1950 ACAGGAATATCTGGTGAAAA	TTTTTCTATTCGCAGCGCTAGA rPheSerileArgSerAlaArg 1670 TGGAGTGCTGGCTCGTATAAAA eGlyValLeuAlaArgIleLys 1730 TTTGATGGCTTCCCAACAGGA ePheAspGlyPheProThrGly 1790 AGGTAATTTCGTGATTATAGCA sGlyAsnPheValileIleAla 1850 CATGGCGATAAATCTTGCGGTT pMetAlaIleAsnLeuAlaVal 1910 GAGCGCAGGTCAAATTGTTGAG tSerAlaGlyGlnIleValGlu

FIG. 1A (3)

	•			
1990		2010		2030
	PTCCGAGTAGAAG		AAACGGTTAGA	GAATCACATTTTA1
LvsGluGluLeuF	heArgValGluG	luAlaGlvG	luThrValArm	GluSerHisPheTyr
-,000	,			oraserurseneryt
2050		2070	•	2090
ATCTGCAGTGATA	GTCAGTATAAGC'	TTAACTTAA	TCGCGAATCAG	ATCCGGTTGCTGAGA
IleCysSerAspS	SerGlnTyrLysLe	euAsnLeuI	leAlaAsnGln	IleArgLeuLeuArg
	• •			
2110		2130		2150
AAAGAAGATCGAG	STAGACGTAATAT:	TTATCGATT	ACTTGCAGTTG	ATCAACTCATCGGTT
LysGluAspArgV	/alAspValIlePl	helleAspI	yrLeuGlnLeu:	IleAsnSerSerVal
. 2120		2100		
2170		2190		2210
GGAGAAAATCGTC	AAAATGAAATAG	CAGATATAT	CTAGAACCTTA	AGAGGTTTAGCCTCA
GIYGIUASNAFGG	inasngiuilea.	Tavabites	erArgThrLew	ArgGlyLeuAlaSer
2230		2250		2270
	` Ր ጥል ጥል ርጥጥጥር ጥጥ'		*****	2270 GTTGAGGATAGAGCA
GlulauAcnila	rollavalcuet.	enserGlet	INICINGANAN	JiTGAGGATAGAGCA ValGluAspArgAla
Olube unsiller	TOTTE VATC 1 3 D.	6026101112	enservidras	va rozuwsby i dy ra
2290		2310		2330
AATAAAGTTCCCA	TGCTTTCAGATT	TGCGAGACA	GCGGTCAAATA	GAGCAAGACGCAGAT
AsnLysValProM	letLeuSerAspLe	euArgAspS	erGlyGlnIle	GluGlnAspAlaAsp
·	•	•		
2350		2370		2390
GTGATTTTGTTTA	\TCAATAGGAAGG/	NA TCGTCTT	CTAATTGTGAG	NTAACTGTTGGGAAA
ValileLeuPhei	leAsnArgLysG	luSerSerS	erAsnCysGlu	lleThrValGlyLys
2412		2422		
2410		2430		2450
AATAGACATGGAT	courtreterr	CGGTATTAC	ATTTCGATCCA	AAATTAGTAAATTC LysileSerLysPhe
Asharduragras	er agrene sers.	ervarrenu	isPneAspPro	razitezetrazue
2470		2490		2510
	AAGTATGGTAAA		CTGCCACTTCAT	CAAAAGTCCTATCC
SerAlaIleLysi	LysValTrpEnd			
ORF2	:>> MetValAsı	nTyrSerAs	nCysHisPheI	leLysSerProIlei
		-	•	
2530		2550		2570
ACCTTGAAAATCA	IGAAGTTTGGAAG	AAGACCTGG	TCAATCTATTA	AGATATCTCCCAAAT
isLeuGluAsnGl	nLysPheGlyAr	gArgProGl	yGlnSerIleLy	slleSerProLysL
2500		2614		
2590		2610		2630
TUGCTCAAAATGG	OATGGTAGAAGT	TATAGGTC	TIGATITICTIT	CATCTCATTACCAT
edy190111V21101	.ynecvalGluva.	Tiregrape	avabluerens	erSerHisTyrHis/
2650		2670		2690
	CCAAAGATTACT		GAATTACAAGG	GAACACAAAAGGGG
laLeuAlaAlaIl	eGlnArgLeuLe	uThrAlaTh	rAsnTvrLvsG	lyAsnThrLysGlyV
	,			-,
2710		2730		2750
TTGTTTTATCCA	SAGAATCAAATAG	TTTTCAATI	TGAAGGATGGA:	PACCAAGAATCCGTT
alValLeuSerA	rgGluSerAsnSe	rPheGlnPh	eGluGlyTrpI	leProArgIleArgP
2770		2790		2810
TTACAAAAACTG	AATTCTTAGAGGC	TTATGGAGI	TAAGCGGTATA	AAACATCCAGAAATA
nerntlystnfG	LUPNELEUGIUAl	スプンとほしいびょ	II LVSAとのTvet。	veThrear arabent

2830 2850 2870 AGTATGAGTTTAGTGGAAAAGAAGCTGAAACTGCTTTAGAAGCCTTATACCATTTAGYSTYrGluPheSerGlyLysGluAlaGluThrAlaLeuGluAlaLeuTyrHisLeuG 2890 2910 2930 ATCAACCGTTTTTAATAGTGGCAACTAGAACTCGATGGACTAATGGAACACAAATAG isGlnProPheLeuIleValAlaThrArgThrArgTrpThrAsnGlyThrGlnIleu 2950 2970 2990 ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAGSARGTTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAGSARGTYrGlnThrLeuSerProIleIleArgIleTyrGluGlyTrpGluGlyLeug 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAG spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAGG leLysProAlaAsnValTyrGlnGluIleLysMetArgPheProAsnAlaSerLysT	
ystyrGluPheSerGlyLysGluAlaGluThrAlaLeuGluAlaLeuTyrHisLeuc 2890 2910 2930 ATCAACCGTTTTTAATAGTGGCAACTAGAACTCGATGGACTAATGGAACACAAATAG isGlnProPheLeuIleValAlaThrArgThrArgTrpThrAsnGlyThrGlnIleu 2950 2970 2990 ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAAS spArgTyrGlnThrLeuSerProIleIleArgIleTyrGluGlyTrpGluGlyLeuc 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAG spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAGG	
ystyrGluPheSerGlyLysGluAlaGluThrAlaLeuGluAlaLeuTyrHisLeuc 2890 2910 2930 ATCAACCGTTTTTAATAGTGGCAACTAGAACTCGATGGACTAATGGAACACAAATAG isGlnProPheLeuIleValAlaThrArgThrArgTrpThrAsnGlyThrGlnIleu 2950 2970 2990 ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAAS spArgTyrGlnThrLeuSerProIleIleArgIleTyrGluGlyTrpGluGlyLeuc 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAG spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAGG	3GA(
ATCAACCGTTTTTAATAGTGGCAACTAGAACTCGATGGACTAATGGAACACAAATAG isGlnProPheLeuileValAlaThrArgThrArgTrpThrAsnGlyThrGlnIlev 2950 2970 2990 ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAA spArgTyrGlnThrLeuSerProIleIleArgIleTyrGluGlyTrpGluGlyLeus 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAG spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	Slyı
2950 2970 2990 ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTACSPATGTYTGINTLeuSerProlleileargileTyrGluGlyTrpGluGlyLeus 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTTAACACCTTTTAATTCACCACCTACACGGAAACSPGluGluAsnileAspileAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTCYSGlyPheValValGluProCysProIleLeuValAspGlnileGluSerTyrPhevilland 1330 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
2950 2970 2990 ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAGSATGTYrGlnThrLeuSerProlleileArgileTyrGluGlyTrpGluGlyLeus 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAC spGluGluAsnileAspileAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnileGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
ACCGTTACCAAACTCTTTCTCCGATCATTAGGATTTACGAAGGATGGGAAGGTTTAL spArgTyrGlnThrLeuSerProIleIleArgIleTyrGluGlyTrpGluGlyLeu 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAC spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	/al/
spArgTyrGlnThrLeuSerProlleIleArgIleTyrGluGlyTrpGluGlyLeus 3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAC spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
3010 3030 3050 ACGAAGAAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAC spGluGluAsnileAspileAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnileGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
ACGAAGAAATATAGATATAGACTTAACACCTTTTAATTCACCACCTACACGGAAAC spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAATGCGTTTCCCAAATGCATCAAAG	lh r i
spGluGluAsnIleAspIleAspLeuThrProPheAsnSerProProThrArgLysi 3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
3070 3090 3110 AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPheV 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
AAGGGTTCGTTGTAGAGCCATGTCCTATCTTGGTAGATCAAATAGAATCCTACTTTC ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPheV 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	lisl
ysGlyPheValValGluProCysProIleLeuValAspGlnIleGluSerTyrPhev 3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
3130 3150 3170 TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	
TCAAGCCTGCAAATGTATACCAAGAAATAAAAATGCGTTTCCCAAATGCATCAAAG	/all
latueDeallakenUalTufGlnGluTlafucMatktaDhaDtnkenklaSarfueT	
Tebysetonianshvariytornordriebyshechtyeheetonshmidsetbys.	[yt]
3190 3210 · 3230	
CTTACACATTTATCGACTGGGTGATTACAGCAGCTGCGAAAAAGAGACGAAAATTA	CT
laTyrThrPheIleAspTrpValIleThrAlaAlaAlaLysLysArgArgLysLeu?	rnrı
3250 3270 3290	
AGGATAATTCTTGGCCAGAAAACTTGTTATTAAACGTTAACGTTAAAAGTCTTGCA	
ysAspAsnSerTrpProGluAsnLeuLeuLeuAsnValAsnValLysSerLeuAla?	ryr
3310 3330 3350	
TTTTAAGGATGAATCGGTACATCTGTACAAGGAACTGGAAAAAAATCGAGTTAGCT	
leLeuArgMetAsnArgTyrIleCysThrArgAsnTrpLysLysIleGluLeuAla	ile
3370 3390 3410	
ATAAATGTATAGAAATCGCCATTCAGCTTGGCTGGTTATCTAGAAGAAAACGCATT	GAA
spLysCysIleGluIleAlaIleGlnLeuGlyTrpLeuSerArgArgLysArgIleG	31 u i
3430 3450 3470	
TTCTGGATTCTTCTAAACTCTCTAAAAAA GAAATTCTATATCTAAATAAAGAGCGC	
heLeuAspSerSerLysLeuSerLysLysGluIleLeuTyrLeuAsnLysGluArg	Phe
3490 3510 3530	
AAGAAATAACTAAGAAATCTAAAGAACAAATGGAACAATTAGAACAAGAATCTATT	
luGluIleThrLysLysSerLysGluGlnMetGluGlnLeuGluGlnGluSerIle	ASN
3550 3570 3590	
AATAGCAAGCTTGAAACTAAAAACCTAATTTATTTAAAGCTCAAAATAAAAAAGAG nd	rtt'

3670 3690 3710 ATCAAAGTTGGGCAAATGACAGGACCGCTCAAGGACCAGCAAATAATCCTTGGGACAACA 11eLysValGlyGlnMetThrGluProLeuLysAspGlnGlnIleIleLeuGlyThrThr

3630

AAAATGGGAAATTCTGGTTTTTATTTGTATAACACTGAAAACTGCGTCTTTGCTGATAAT ORF3>> MetGlyAsnSerGlyPheTyrLeuTyrAsnThrGluAsnCysValPheAlaAspAsn

3650

		FIG. 14 (5)
3730	3750	3770
TCAACACCTGTCGCAGCC	AAAATGACAGCTTCTGAT	GGAATATCTTTAACAGTCTCCAAT
SetintProvatataAta	LysmetIntAlaSerAsp	GlyIleSerLeuThrValSerAsn
3790	3810	3830
AATTCATCAACCAATGCT	TCTATTACAATTGGTTTG	GATGCGGAAAAAGCTTACCACCTT
AsnSerSerThrAsnAla	SerIleThrIleGlyLeu	AspAlaGluLysAlaTyrGlnLeu
	·	·
3850	3870	3890
TlafauGlutvetauGlu	AspGlatletenkspGl	ATTGCTGATACTATTGTTGATAGT IleAlaAspThrIleValAspSer
reconditional	vaborutterenvabotå	11eAlaAspInfileValAspSer
3910	3930	3950
ACAGTCCAAGATATTTTA	GACAAAATCAAAACAGAC	CCTTCTCTAGGTTTGTTGAAAGCT
ThrValGlnAspileLeu	AspLysIleLysThrAsp	ProSerLeuGlyLeuLeuLysAla
3970	3990	4010
TTTAACAACTTTCCAATC	ACTAATAAAATTCAATGC	AACGGGTTATTCACTCCACTAAC
PheAsnAsnPheProIle	ThrAsnLysIleGlnCys.	AsnGlyLeuPheThrProSerAsn
4030	4050	4070
		4070 TTCACAGTCACACCCAAAAGCTCT
IleGluThrLeuLeuGly	GlyThrGluIleGlyLys	PheThrValThrProLysSerSer
•		
4090	4110	4130 ·
GGGAGCATGTTCTTAGTC	TCAGCAGATATTATTGCA	TCAAGAATGGAAGGCGCGTTGTT
GlySet MecPheLeuval	serataaspiteiteata:	SerArgMetGluGlyGlyValVal
4150	4170	4190
CTAGCTTTGGTACGAGAA	GGTGATTCTAAGCCCTGC	GCGATTAGTTATGGATACTCATCA
LeuAlaLeuValArgGlu	GlyAspSerLysProCys	AlalleSerTyrGlyTyrSerSer
4210	4230	4250
	AGTCTAAGAACCAGTATT	4250 ACTAATACAGGATTGACTCCGACA
GlyIleProAsnLeuCys	SerLeuArgThrSerIle'	ThrAsnThrGlyLeuThrProThr
4270		
	4290	4310
ThrTvrSerLeuargVal	GCGGTTTAGAAAGCGGT(GlyGlyLauGlySa=Cl++	GTGGTATGGGTTAATGCCCTTTCT ValValTrpValAsnAlaLeuSer
The Lytoet beant grant	orlorl renot no et Gill	valvallipvalAshAlaLeuSer
4330	4350	4370
AATGGCAATGATATTTTA	ggaataa caaatacttct.	AATGTATCTTTTTTAGAGGTAATA
AsnGlyAsnAspIleLeu	GlyIleThrAsnThrSer	AsnValSerPheLeuGluValIle
4390	4410	4430
	ACAATTTTTATTGGATTT	TICTTATAGGTTTTATATTTAGAG
ProGlnThrAsnAlaEnd		
4450	4470	
	4470 GGGTTTGTTATCCAAAA	4490 AAAAGAAAAGTGAGGGACGATTTT
	ORF4 >> MetGlnAcni	LysArgLysValArgAspAspPhe
		-1 3-1 a ratur Awahwabhue
4510	4530	4550
ATTAAAATTGTTAAAGAT	GTGAAAAAAGATTTCCCCC	GAATTAGACCTAAAAATACGAGTA
rierAzite satrAzyzb	valLysLysAspPhePro(GluLeuAspLeuLysIleArgVal
4570	4590	4610
AACAAGGAAAAGTAACT	TTCTTAAATTCTCCCTTA	3 A C T C T A C C A T A A A C T C T C T C
AsnLysGluLysValThr	PheLeuAsnSerProLeu(GluLeuTyrHisLysSerValSer

FIG. 1A (6)

	4630	4650	4670
			GGATTATTCCCAGACTCTCCT
	LeuileLeuGlyLeuLeuG	inginilegluAsnSerLeu	GlyLeuPheProAspSerPro
	4690	4710	4730
	GTTCTTGAAAATTAGAGG	SATAACAGTTTAAAGCTAAAA	AAGGCTTTGATTATGCTTATC
			LysAlaLeuIleMetLeuIle
	4750	4770	4790
			ACTTACTCTAACGTTGGAGTT
•	LeuSerArgLysAspMetI		
	4810	4830	4850
	GATTTGCACACCTTAGTTT	TTTTGCTCTTTTAAGGGAGGA	ACTGGAAAAACAACACTTTCT
ORF5			ThrGlyLysThrThrLeuSer
	4870	4890	4910
			AAAGTGTTACTTGCTGACCTA
	LeuAsnValGlyCysAsnI	LeuAlaGlnPheLeuGlyLys	LysValLeuLeuAlaAspLeu
	4930	4950	4970
			GTCAGAAGTGACCAAAAAGGC
	AspProGlnSerAsnLeus	SerSerGlyLeuGlyAlaSer	ValArgSerAspGlnLysGly
	4990	5010	5030
	TTGCACGACATAGTATAC	<u>ACATCAAACGATTTAAAATCA</u>	ATCATTTGCGAAACAAAAA
	LeuHisAsplleValTyr:	ChrSerAsnAspLeuLysSer	IleIleCysGluThrLysLys
	5050	5070	5090
			CAGTTTAGAGAATTGGATATT
	AspSerValAspLeuIle	ProAlaSerPheSerSerGlu	GlnPheArgGluLeuAspIle
	5110	5130	5150
			GAGTACTGCGCTCCTTTTAT
	HisArgGlyProSerAsn	AsnLeuLysLeuPheLeuAsn	GluTyrCysAlaProPheTyr
	5170	5190	5210
			STTAACGAAAGAAGCTTTTGTT
	AspileCysileIleAsp	ThrProProSerLeuGlyGly	LeuThrLysGluAlaPheVal
	5230	5250	5270
			PTTTTCTATTCTAGGGTTACAA
	AlaGlyAspLysLeuIle	AlaCysLeuThrProGluPro	PheSerIleL euGly LeuGln
	5290	5310	5330
			AGAAGAACACATTCTTGGAATA
	LysileArgGluPheLeu	SerSerValGlyLysProGli	uGluGluHisIleLeuGlyIle
	5350	5370	5390
			AATGTATATAGACATTATCGAG
•	AlaceuserPhetrpAsp	AspargasnserthrasnGli	nMetTyrIleAspIleIleGlu
	5410	5430	5450
	TCTATTTACAAAAACAAG	CTTTTTCAACAAAATTCG	TCGAGATATTTCTCTCAGCCGT
	SerileTyrLysAsnLys	LeuPheSerThrLysIleAr	gArgAspIleSerLeuSerArg
	5470	5490	5510
	TCTCTTCTTAAAGAAGAT	TCTGTAGCTAATGTCTATCC	AAATTCTAGGGCCGCAGAAGAT
	SerLeuLeuLysGluAsp	SerValAlaAsnValTyrPr	oAsnSerArgAlaAlaGluAsp

		FIG.	1 A	(7)	
5530	5550	557	70		
ATTCTGAAGTTAACGCATGAAATAG	<u>CAAATATTTTGCATATCGA</u>	ATAT	BAAC	GAGATTA	١C
IleLeuLysLeuThrHisGluIleA	laAsnIleLeuHisIleGl	uTyr	SluA	rgAspTy	ŢĽ
5590	5610	563	30		
TCTCAGAGGACAACGTGAACAAACT	rarara grageggatgte	TTTTI	AATT	AAAAAA 1	C
SerGlnArgThrThrEnd			_	_	
ORF6 >> ValAsnLysLet	urystysGluAlaAspVal	PhePi	ery	sLysAsn	ıG
5650	5670	569	0		
AAACTGCCGCTTCTCTAGATTTTAA	GAAGACGCTTCCCTCCATT	GAACI	TTA	CTCAGCA	Ŋ
InThrAlaAlaSerLeuAspPheLys	stystnrteuproserile	GluLe	uPh	eSerAla	T
5710	5730	575	0		
CTTTGAATTCTGAGGAAAGTCAGAG	PTTGGATCGATTATTTTTA	TCAGA	GTC	CCAAAAC	T
hrLeuAsnSerGluGluSerGlnSer	rLeuAspArgLeuPheLeu	SerGl	uSe	rGlnAsn	T
5770	5790	581	.0		
ATTCGGATGAAGAATTTTATCAAGAA	^a gacatcctagcggtaaaa	CTGC1	TAC:	IGGTCA G	À
yrSerAspGluGluPheTyrGlnGlu	uAspIleLeuAlaValLys	LeuLe	uTh	rGlyGln	ιI
5830	5850	587			
TAAAATCCATACAGAAGCAACACGT	ACTTCTTTTAGGAGAAAA	ATCTA	TAA!	rgc taga	A
leLysSerIleGlnLysGlnHisVa	lLeuLeuLeuGlyGluLys	IleTy	rAsı	nAlaArg	L
5890	5910	593	0		
AAATCCTGAGTAAGGATCACTTCTCC	CTCAACAACTTTTTCATCT	TGGAT	'AGA	STTAGTT	T
ysIleLeuSerLysAspHisPheSe	rSerThrThrPheSerSer	TrpIl	eGlu	ıLeuVal	P
5950	5970	599	0		
TTAGAACTAAGTCTTCTGCTTACAA	rgctcttgcatattacgag	CTTTT	'TAT'	MACCIC	C
heArgThrLysSerSerAlaTyrAsi	uwranenwraithithicin	LeuPh	elle	eAsnLeu	P
6010	6030	605	0		
CCAACCAAACTCTACAAAAAGAGTT	ICAATCGATCCCCTATAAA	TCCGC	ATA	PATTTTG	G
roAsnGlnThrLeuGlnLysGluPho	ecinserileprotyrLys	Seral	ату	IleLeu	À
6070	6090	611	.0		
CCGCTAGAAAAGGCGATTTAAAAACC	CAAGGTCGATGTGATAGGG	AAAGT	'ATG	GGAATG	T
laAlaArgLysGlyAspLeuLysTh	crassarvabsarriecta	rysva	TCA	GlyMet	S
6130	6150	617	_		
CGAACTCATCGGCGATAAGGGTGTT	GGATCAATTTCTTCCTTCA	TCTAG	AAA	CAAAGAC	G
erAsnSerSerAlaIleArgValLe	uAspGlnPheLeuProSer	SerAr	gAsı	ıLysAsp	V
6190	6210	623	0		
TTAGAGAAACGATAGATAAGTCTGA	PTCAGAGAAGAATCGCCAA	TTATC	TGAT	TTTCTTA	A
alArgGluThrIleAspLysSerAs	pserGluLysAsnArgGln	LeuSe	rAsı	PheLeu	I
6250	6270	629	0		
TAGAGATACTTCGCATCATGTGTTCC	CGGAGTTTCTTTGTCCTCC	TATAA	CGA	LAATCTT	C
leGluIleLeuArgIleMetCysSe	retyvatserLeuSerSer	TYTAS	nGl	AsnLeu	L
6310	6330	635	0		
TACAACAGCTTTTTGAACTTTTTAAC	GCAAAAGAGCTGATCCTCC	GTCAG	CTC	ATATATA	T
euGlnGlnLeuPheGluLeuPheLy	sGlnLysSerEnd				

FIG. 1A (8)

6370	6390	6410
ATATCTATTATATAT	^r atatattagggatttgatttc	ACGAGAGAGATTTGCAACTCTTG
6430	6450	6470
GTGGTAGACTTTGC		ACTCTTGGTGGTAGACTTTGCAA
6490	6510 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	6530 Aaaaatettattattaaa
		AND THE PARK TO TAKE
6550	6570	6590
GCTCCGATTTTGAAT	PAGCTTTGGTTAAGAAAATGGGC	TCGATGGCTTTCCATAAAGTAG SerHetAlaPheHisLysSerAr
OMI /	>> bedvalbyabyametdly	Serue CYTS NEHI STAR SELVE
6610	6630	6650
ATTGTTTTTAACTTT	TGGGGACGCGTCGGAAATTTGG	TTATCTACTTTATCTATCTAAC LeuSerThrLeuSerTyrLeuTh
gueurneneurntri.	redryksparaserGruffeTfp	preuserint Leuser Tyr LeuTh
6670	6690	6710
TAGAAAAAATTATGO	GTCTGGGATTAACTTTCTTGTT	TCTTTAGAGATTCTGGATTTATC
twidnaswautatwi	.aserGIYIIEASHPHELEUVAI	SerLeuGluIleLeuAspleuSe
6730	6750	6770
GGAAACCTTGATAAA	GGCTATTTCTCTTGACCACAGC	GAATCTTTGTTTAAAATCAAGTC
reignurregitera	/salalleSerLeuAspHisSer	GluSerLeuPheLysIleLysSe
6790	6810	6830
TCTAGATGTTTTAA	TGGAAAAGTTGTTTCAGAGGCA	TCTAAACAGGCTAGAGCGGCATG
rLeuAspvalPneAs	inGiyLysValValSerGluAla	SerLysGlnAlaArgAlaAlaCy
6850	6870	6890
CTACATATCTTTCAC	AAAGTTTTTGTATAGATTGACC	AAGGGATATATTAAACCCGCTAT
sryttiesetrnern	irLysPheLeuTyrArgLeuThr	LysGlyTyrIleLysProAlaIl
6910	6930	6950
TCCATTGAAAGATTT	TGGAAACACTACATTTTTAAA	ATCCGAGACAAAATCAAAACAGA
erroceutysaspro	iegryksnintinthnephelys	IleArgAspLysIleLysThrGl
6970	6990	7010
ATCGATTTCTAAGCA	GGAATGGACAGTTTTTTTGAA	GCGCTCCGGATAGTGAATTATAG
nzetttezetråget	InGlutrpThrValPhePheGlu	AlaLeuArgIleValAsnTyrAr
7030	7050	7070
AGACTATTTAATCG	TAAATTGATTGTACAAGGGATC	CGTAAGTTAGACGAAATTTTGTC
gAspTyrLeu11eG1	lyLysLeuIleValGlnGlyIle	ArgLysLeuAspGluIleLeuSe
7090	7110	7130
TTTGCGCACAGACGA	\TCTATTTTTTGCATCCAATCAG	ATTTCCTTTCGCATTAAAAAAA
rLeuArgThrAspAs	spLeuPhePheAlaSerAsnGln	IleSerPheArgIleLysLysAr
7150	7170	7190
ACAGAATAAAGAAA	CCAAAATTCTAATCACATTTCCT	'ATCAGCTTAATGGAAGAGTTGCA
gGlnAsnLysGluTh	arLysIleLeuIleThrPhePro	lleSerLeuMetGluGluLeuGl
7210	7230	7250
AAAATACACTTGTG	GAGAAATGGGAGAGTATTTGTT	TCTAAAA TAGGGA TTCCTGTAAC
nLysTyrThrCysG]	lyArgAsnGlyArgValPheVal	SerLysIleGlyIleProValTh

FIG. 1A (9)

7270 7290 7310
AACAAGTCAGGTTGCGCATAATTTTAGGCTTGCAGAGTTCCATAGTGCTATGAAAATAAA
rThrSerGlnValAlaHisAsnPheArgLeuAlaGluPheHisSerAlaMetLysIleLy

7330 7350 7370
AATTACTCCCAGAGTACTTCGTGCAAGCGCTTTGATTCATTTAAAGCAAATAGGATTAAA
sileThrProArgValLeuArgAlaSerAlaLeuIleHisLeuLysGlnIleGlyLeuLy

7390 7410 7430
AGATGAGGAAATCATGCGTATTTCCTGTCTTTCATCGAGACAAAGTGTGTTCTTATTG
sAspGluGluIleMetArgIleSerCysLeuSerSerArgGlnSerValCysSerTyrCy

7450 7470 7490
TTCTGGGGAAGAGGTAATTCCTCTAGTACAAACACCCCACAATATTGTGATATAATTAAAA
sSerGlyGluGluValIleProLeuValGlnThrProThrIleLeuEnd

TT

FIG. 1B (1)

ATGAATCTGCAAGAGTTTTCCCTTTTAAATCATAATCTAATCCTAGATGCGATTAAAAA MetAsnLeuGlnGluPheSerLeuLeuAsnHisAsnLeuIleLeuAspAlaIleLysLys

GTTTCCTCTGCCAAGACTTCTTGGACCGAAGGTACTAAACAAGTTCGAGCAGCCAAGCTAT ValSerSerAlaLysThrSerTrpThrGluGlyThrLysGlnValArgAlaAlaSerTyr

ATTTCCTTAACAAGATTCCTAAACAGGATGACTCAAGGAATAGTCGCTATAGCGCAACCT IleSerLeuThrArgPheLeuAsnArgMetThrGlnGlyIleValAlaIleAlaGlnPro

TCTARACAAGAAATAGTCGAACATTTTTTAAAACCAGGGAAATAGTAAAAACGGATGCGSerLysGlnGluAsnSerArgThrPhePheLysThrArgGluIleValLysThrAspAla

TGGTTGATCGCCCAGACAATGCTCCAAGGAGGTAAACGCTCCTCTGAAGTCTTAAGCTTG TrpLeulleAlaGlnThrMetLeuGlnGlyGlyLysArgSerSerGluValLeuSerLeu

GAGATTAGTCAGATTTGTTTCCAACAAGCTACCATTTCTTCTCCCAGCTTAAGAACCGT GluileSerGlnileCysPheGlnGlnAlaThrIleSerPheSerGlnLeuLysAsnArg

CAGACAGAAAAGAGGATTATTATAACTTATCCTCAGAAGTTTATGCACTTTCTACAAGAGGInThrGluLysArgIleIleIleThrTyrProGlnLysPheMetHisPheLeuGlnGlu

FIG. 1B (2)

TACATCGGTCAACGAAGAGGTTTTGTCTTCGTAACTCGCTCCGGAAAAATGGTGGGGTTA TyrlleGlyGlnArgArgGlyPheValPheValThrArgSerGlyLysMetValGlyLeu

AGGCAAATCGCCCGCACGTTCTCTCAAGCAGGACTACAAGCTGCAATCCCTTTTAAAATA ArgGlnIleAlaArgThrPheSerGlnAlaGlyLeuGlnAlaAlaIleProPheLysIle

ACCCCGCACGTGCTTCGAGCAACCGCTGTGACGGAGTACAAACGCCTAGGGTGCTCAGAC ThrProHisValLeuArgAlaThrAlaValThrGluTyrLysArgLeuGlyCysSerAsp

TCCGACATAATGAAGGTCACAGGACACGCAACCGCAAAGATGATATTTGCGTACGATAAA SerAspIleMetLysValThrGlyHisAlaThrAlaLysMetIlePheAlaTyrAspLys

TCTTCTCGAGAAGACAACGCTTCAAAGAAGATGGCTCTAATATAGCCTAAAGGTGTTTTT SerSerArgGluAspAsnAlaSerLysLysMetAlaLeuIleEnd

TCTGGCAACAGAATATGAATAT

FIG. 2

		F19. 2	
	3610	3630	3650
TAAAA	GGGAAATTCTGGTTTTTATT	TGTATAACACTGAAAACTG	ひとひと でんんかでいるでである。
ORF3>> Me	tGlyAsnSerGlyPheTyrL	BUTVrAsnThrGluAsnCv	:Val Phallalenlen
	•		, vari nenzanaphan
	3670	3690	3710
ATCAA	AGTTGGGCAAATGACAGAGC	CGCTCAAGGACCAGCAAATI	ATCCTTGGGACAACA
IleLy	sValGlyGlnMetThrGluP	roLeuLysAspGlnGlnIle	IleLeuGlyThrThr
man n a 1	3730	3750	3770
TCAACA	ACCTGTCGCAGCCAAAATGA	CAGCTTCTGATGGAATATCT	TTAACAGTCTCCAAT
2611111	rProValAlaAlaLysMetT	utwraseraspGrAtreset	LeurnivalSerAsn
	3790	3810	3830
ልል ጥጥር ያ	ATCAACCAATGCTTCTATTA		AAACCTTACCACCTT
AsnSe	SerThrAsnAlaSerIleT	hrTleGlvLeuAenAleGlu	Luckla Turclata
		oct j so salopataota	-levialledined
	3850	3870	3890
ATTCT	\GAAAAGTTGGGAGATCAAA	TTCTTGATGGAATTGCTGAT	ACTATTGTTGATAGT
IleLev	GluLysLeuGlyAspGlnI	leLeuAspGlyIleAlaAsp	ThrileValAspSer
		_	•
	3910	3930	3950
ACAGT	CAAGATATTTTAGACAAAA	TCAAAACAGACCCTTCTCTA	GGTTTGTTGAAAGCT
ThrVal	GlnAspIleLeuAspLysI	leLysThrAspProSerLeu	GlyLeuLeuLysAla
	3070	3000	
	3970	3990	4010
TTTAAC	CAACTTTCCAATCACTAATA	AAATTCAATGCAACGGGTTA	TTCACTCCCAGTAAC
PHEASI	AsnPheProIleThrAsnLy	AgileGiucAswauGiAren	PhethrProSerAsn
	4030	4050	4070
ATTGAZ	ACTTTATTAGGAGGAACTG		1070 12020000000000000000000000000000000
IleGlu	ThrLeuLeuGlyGlyThrG	luIleGlvI.vcPheThrVal	Thr Brotus Carcar
			onleadrage
	4090	4110	4130
GGGAGG	CATGTTCTTAGTCTCAGCAG	ATATTATTGCAT <mark>CAAGAA</mark> TG	GAAGGCGGCGTTGTT
GlySer	MetPheLeuValSerAlaA	spileileAlaSerArgMet	GluGlyGlyValVal.
	4150	4170	4190
CTAGCT	TTGGTACGAGAAGGTGATT	CTAAGCCCTGCGCGATTAGT	TATGGATACTCATCA
LeuAls	LeuValArgGluGlyAspS	erLysProCysAlaIleSer	TyrGlyTyrSerSer
	4210	4230	4250
GGCATI	CCTAATTTATGTAGTCTAA		425U
GlvIle	ProAsnLeuCysSerLeuA	raThrSerTleThrAenThr	Clutauthereathe
			orgreaturerorur
•	4270	4290	4310
ACGTAT	TCATTACGTGTAGGCGGTT	TAGAAAGCGGTGTGGTATGG	GTTAATGCCCTTTCT
ThrTyr	SerLeuArgValGlyGlyL	euGluSerGlyValValTrp	ValAsnAlaLeuSer
		•	
	4330	4350	4370
	CAATGATATTTTAGGAATAA		
AsnGly	AsnAspIleLeuGlyIleT	hrAsnThrSerAsnValSer	PheLeuGIuValIle
	4300	4410	4430
CCTCXX	4390 A C A A A C C C T T A A C A A T T T T T	4410	4430
Profile	ACAAACGCTTAAACAATTT' ThrAsnAlaEnd	I IA I TGGATTTTTCTTATAG	GTTTTATATTTAGAG
1.0011	noiMT&DIM		

FIG. 3

-106 MetSerLysThrThrLysLysPheAsnSerLeuCysIleAspLeuProArgAspLeuSer LeuGluIleTyrGlnSerIleAlaSerValAlaThrGlySerGlyAspProHisSerAsp AspPheThrAlaIleAlaTyrLeuArgAspGluLeuLeuThrLysHisProThrLeuGly SerGlyAsnAspGluAlaThrArgArgThrLeuAlaIleAlaLysLeuArgGluAlaAsn GlyAspArgGlyGlnIleAsnArgGluGlyPheLeuHisAspLysSerLeuSerTrpAsp IleArgAlaThrGlySerMetGlyAsnSerGlyPheTyrLeuTyrAsnThrGluAsnCys ValPheAlaAspAsnIleLysValGlyGlnMetThrGluProLeuLysAspGlnGlnIle IleLeuGlyThrThrSerThrProValAlaAlaLysMetThrAlaSerAspGlyIleSer LeuThrValSerAsnAsnSerSerThrAsnAlaSerIleThrIleGlyLeuAspAlaGlu LysAlaTyrGlnLeuIleLeuGluLysLeuGlyAspGlnIleLeuAspGlyIleAlaAsp ThrIleValAspSerThrValGlnAspIleLeuAspLysIleLysThrAspProSerLeu GlyLeuLeuLysAlaPheAsnAsnPheProIleThrAsnLysIleGlnCysAsnGlyLeu PheThrProSerAsnIleGluThrLeuLeuGlyGlyThrGluIleGlyLysPheThrVal ThrProLysSerSerGlySerMetPheLeuValSerAlaAspIleIleAlaSerArgMet GluGlyGlyValValLeuAlaLeuValArgGluGlyAspSerLysProCysAlaIleSer TyrGlyTyrSerSerGlyIleProAsnLeuCysSerLeuArgThrSerIleThrAsnThr GlyLeuThrProThrThrTyrSerLeuArgValGlyGlyLeuGluSerGlyValValTrp ValAsnAlaLeuSerAsnGlyAsnAspIleLeuGlyIleThrAsnThrSerAsnValSer PheLeuGluValIleProGlnThrAsnAlaEnd



EUROPEAN SEARCH REPORT

EP 91 10 6110

	DOCUMENTS CONSIDERE		1	<u> </u>
Category	Citation of document with indication, of relevant passages	where appropriate,	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
x	M. COMANDUCCI ET AL. 'Identif	ication and	1-3,6,	C12N15/30
	characterization of a 28-kd pr	rotein of C.	14,20,24	A61K39/118
	trachomatis encoded by the 7.5	i-kb common		G01N33/569
.	plasmid'		i	C07K15/00
	1990 . IN 'CHLAMYDIA INFECTION	IS', ED, W.R. BOWIE		
	ET AL. , CAMBRIDGE UNIVERSITY	PRESS, NY, US		
	pages 121 - 124		1	
	* Whole article *			
- 1				
×	EP-A-0 336 412 (SCLAVO S.P.A.)	1	1-19	
	* Whole document *			
r			20-33	1
.				1
Y	INFECTION AND IMMUNITY		20-33	1
l	vol. 57, no. 9, September 1989	, us	1	1
	pages 2683 - 2690;		İ	
l	S. BIRKELUND ET AL: 'Character			1
ŀ	and recombinant 75-kilodalton	_		İ
- 1	Chlamydia-trachomatis serovar	L2'		1
- 1	* Whole article *	•	·	·
,	PLASMID		1-19	TECHNICAL FIELDS SEARCHED (Int. Cl.5)
•	vol. 23, 1990, NEW YORK, US		1-19	
- 1	pages 149 - 154;			C12N
Į	M. COMANDUCCI ET AL.: 'Diversi	laur of Abo		A61K
1	Chlamydia trachomatis common p	_		C07K
- 1	with different pathogenicity'	INTERIOR IN DIGARIZ		CO. K
- 1	* Whole article *			1 ·
i				1
.	MOLECULAR MICROSIOLOGY		1-19	
	vol. 2, no. 4, 1988, OXFORD, G	28		
	pages 531 - 538;			1
Į	M. COMANDUCCI ET AL.: 'The str	ucture of a	Ţ	1
	plasmid of Chlamydia trachgmat			1
i	required for growth within man			1
i	* Whole article *			1
r I			20-33	1
	***			1
- 1		-/		
			1	1.
			J	
	The present search report has been drawn	up for all claims		
	Place of south	Date of completion of the search	1-1	Econom
	BERLIN	19 MAY 1992) JUL	JA P.
	ATTOON OF CITED BOOK AS		1 1 - 1 - 1	
•	ATEGORY OF CITED DOCUMENTS	T : theory or princi E : cariler patent é		
	icularly relevant if taken alone	after the filing	date	
s : juiro	icularly relevant if combined with another ment of the same category	D : document citéé L : document citéé		
A : tech	nological background	***************************************		
O : con-written disclosure P : intermediate document			serae bereen 1939	112, conscionaring

N 1500 CC.EZ (POMOL)

EUROPEAN SEARCH REPORT

EP 91 10 6110 PAGE2

DOCUMENTS CONSIDERED TO BE RELEVANT					PAGE2	
Category	Citation of document with of relevant p	indication, where appropriate,		Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. CL5)	
Y	INFECTION AND IMMUNITY vol. 59, no. 1, Januar pages 79 - 90; M.C. CERRONE ET AL.: the gene for heat shoc Chlamydia trachomatis reactivity of the prot * Whole article *	y 1991, US Cloning and sequence of k protein 60 from and immunological	20	1-33		
×		LONDON, GB sis of the entire the cryptic plasmid of serovar L1. Evidence for	1-	-19		
Y		<u>.</u>	20	∸33		
Υ	EP-A-0 293 079 (368800 * Whole document *	ALBERTA LTD)	20	⊢33	TECHNICAL FIELDS SEARCHED (bs. Cl.5)	
X	NUCLEIC ACIDS RESEARCH vol. 15, no. 24, 1987, page 10596; K.S. SRIPRAKASH ET AL. protein in chlamydial * Whole article *	LONDON, GB : 'A gene for dnaB like	1-	4, 12		
K	• •	: 'Characterization and from the trachoma biovar	1-	19		
Υ .		-/	20-	-33		
	The present search report has I	een drawn up for all claims	\dashv			
	Place of search BERLIN	Date of completion of the search 19 MAY 1992		JULI	Branine A P.	
X : part Y : part docs A : tech O : non-	CATEGORY OF CITED DOCUME icalarly relevant if taken alone culsarly relevant if combined with an meet of the same category nological background written disclosure mediate document	E : earlier patent after the filin	docume g date ed in the ed for oti	nt, but publi application ber reasons	ished on, or	



EUROPEAN SEARCH REPORT

Application Number

EP 91 10 6110

	DOCUMENTS CONSIDEI	PAGE3		
Category	Citation of document with indicate of relevant passages	ion, where appropriate,	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. CL5)
,	EP-A-0 192 033 (CHIRON CORP	ORATION)	20-33	
1	* Whole document *	·	i	
ı				
- 1		•		
			[
i		·	į	
l			1	
- 1				
			,	
l				
j				
- 1				
l			}	
l				
l				
l			<u> </u>	
- 1				
ŀ			l l	·
1				
- 1				•
- 1				
i				TECHNICAL PIELDS
				SEARCHED (Int. CL5)
- 1				
l l			1	
			1 1	
- 1				
1				
1				
			1	
i		`	ł	
		•	i	
			Ì	
	•		1	
				•
			1	
		•		
			i i	
			•	
	The assessed county bear have the	ann na fao all abrino	-	
	The present search report has been dr	Dub of completion of the search		Braine
	BERLIN	19 MAY 1992	JULI	A P.
•	CATEGORY OF CITED DOCUMENTS	T : theory or pr	inciple underlying the it document, but publi	invention shed on, or
X : part	icularly relevant if taken alone	after the fill	ne date	
Y : part	cularly relevant if combined with another	D : éocazaent d	ted in the application	
doc:	ment of the same category	L : document ci	ted for other reasons	
	aological background	A · manker of t	he came potent femile	. corresponding
O : non-written disclosure & : member of the same pate				

KPO FORM LSC CLEZ (POICE)